

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:46:24 ; Search time 86.5929 Seconds
(without alignments)
2086.045 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MAGIAAKAKREARAGLS.....PENTGTIEDLISWLCFSVL 700

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhcc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	99.3	700	6 Q9GLG1	Q9GLG1 macaca fasc
2	2568.5	69.2	705	13 Q9YIC1	Q9YIC1 colutrinx co
3	2480	66.8	724	13 Q8UW96	Q8UW96 xenopus lae
4	2464.5	66.4	714	6 Q9GLG2	Q9GLG2 macaca fasc
5	2463	66.3	715	13 Q42133	Q42133 gallus gall
6	2455.5	66.1	716	6 Q9N185	Q9N185 bos taurus
7	2453.5	66.1	716	6 Q9N0V6	Q9N0V6 bos taurus
8	2426	65.3	713	11 Q8C2J1	Q8C2J1 mus musculu
9	2381	64.1	704	13 Q918G2	Q918G2 brachydanto
10	2356.5	63.5	703	11 Q84698	Q84698 rattus norv
11	2335	62.9	530	11 Q8BVP9	Q8BVP9 mus musculu
12	2322.5	62.6	703	11 Q91VA3	Q91VA3 mus musculu
13	2309	62.2	702	13 Q918T0	Q918T0 xenopus lae
14	2279.5	61.4	650	6 Q9N0M7	Q9N0M7 sus scrofa
15	2142	57.7	421	11 Q8R486	Q8R486 mus musculu
16	2014	54.2	702	4 Q8N4R5	Q8N4R5 homo sapien

17	1959.5	52.8	709	6 Q9XSJ3	Q9XSJ3 oryctolagus
18	1954.5	52.6	709	6 Q9XSJ1	Q9XSJ1 bos taurus
19	1953.5	52.6	709	6 Q9XSJ2	Q9XSJ2 sus scrofa
20	1952.5	52.6	709	11 Q08702	Q08702 rattus norv
21	1946.5	52.4	701	13 Q8AYE8	Q8AYE8 xenopus lae
22	1943.5	52.3	709	11 Q8B977	Q8B977 mus musculu
23	1934.5	52.1	737	11 Q70376	Q70376 rattus norv
24	1925.5	51.9	757	11 Q70482	Q70482 rattus norv
25	1905.5	51.3	755	11 Q9QZP9	Q9QZP9 rattus norv
26	1882	50.7	690	4 Q14815	Q14815 homo sapien
27	1875.5	50.5	821	6 Q46596	Q46596 sus scrofa
28	1874.5	50.5	822	6 Q9TTH8	Q9TTH8 bos taurus
29	1867.5	50.3	822	6 Q9TTH9	Q9TTH9 ovis aries
30	1858.5	50.1	815	6 Q9GLG7	Q9GLG7 macaca fasc
31	1847	49.7	690	11 Q9D805	Q9D805 mus musculu
32	1844	49.7	688	11 Q35919	Q35919 mus musculu
33	1841	49.6	674	11 Q35920	Q35920 rattus norv
34	1795	48.3	664	4 Q9NS74	Q9NS74 homo sapien
35	1610.5	43.4	925	5 Q9VT65	Q9VT65 drosophila
36	1592.5	42.9	925	5 Q96454	Q96454 drosophila
37	1521	41.0	381	11 Q91UZ9	Q91UZ9 mus musculu
38	1453	39.1	720	11 Q9ER56	Q9ER56 mus musculu
39	1354	36.5	565	5 Q9V8U6	Q9V8U6 drosophila
40	1291	34.8	493	11 Q9R113	Q9R113 rattus norv
41	1213.5	32.7	760	5 Q45033	Q45033 schistosoma
42	1207.5	32.5	758	5 Q96072	Q96072 schistosoma
43	1194.5	32.2	758	5 Q96071	Q96071 schistosoma
44	1161	31.3	575	5 Q8MQV0	Q8MQV0 homarus ame
45	1120.5	30.2	462	11 Q9ER54	Q9ER54 mus musculu

ALIGNMENTS

RESULT 1

Q9GLG1 ID Q9GLG1 PRELIMINARY; PRT; 700 AA.
AC Q9GLG1;
DT 01-MAR-2001 (T-REMBLrel. 16, Created)
DT 01-MAR-2001 (T-REMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Calpain 2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Nakajima T., Fukiage C., Azuma M., Shearer T.R.;
RT "Calpain isoforms in the eye of monkey."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF284441; AAG22771.1; -.
DR HSSP; P04574; IALV.
DR MEROPS; C02.002; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease_C2.
DR InterPro; IPR000169; SHPoc_acsite.
DR Pfam; PF01067; Calpain III; 1.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00648; Peptidase C2; 1.
DR PRINTS; PR00704; CALPAIN_III; 1.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; CyfPc; 1.
DR SMART; SM00054; Efn; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
SQ SEQUENCE 700 AA; 80026 MW; DCEH16214F05057C CRC64;

Query Match 99.3%; Score 3687; DB 6; Length 700;
Best Local Similarity 99.1%; Pred. No. 7, 5e-272;
Matches 694; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 MAGIAAKLAKDREAEGLSHERAIKYNODYEALRNECLAEAGTLFODPSFPAISALGF 60
Db 1 MAGIAAKLVKDRBAEAGLSHERAIKYNODYEALRNECLAEAGTLFODPSFPAISALGF 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Db 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Qy 121 LARVPLNOSFOENAGIFHFQFQWQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 180
Db 121 LARVPLNOSFOENAGIFHFQFQWQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIQXALQKSLGCG 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIQXALQKSLGCG 240
Qy 241 SIDITSAADSEALITFOKLVKHAIVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDN 300
Db 241 SIDITSAADSEALITFOKLVKHAIVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDN 300
Qy 301 CPSNNTIDPEERERLTRHEDGEFMSFSDFLRHYSRLICNLTPDILTSDTYKKMKLT 360
Db 301 CPSNNTIDPEERERLTRHEDGEFMSFSDFLRHYSRLICNLTPDILTSDTYKKMKLT 360
Qy 361 MDGMMRGSTAGGCNRYNPTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRRRQR 420
Db 361 MDGMMRGSTAGGCNRYNPTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRRRQR 420
Qy 421 KMGEDMHTIGFYIYVEPEELSGQTNILSKNPFLLNRAERSDFTINREVLNRPKLPG 480
Db 421 KMGEDMHTIGFYIYVEPEELSGQTNILSKNPFLLNRAERSDFTINREVLNRPKLPG 480
Qy 481 EYIIVPSTFEENKGDPCIRVPESEKKADYQAVDDEIENLEBPISDIDDDGVRRLPAQ 540
Db 481 EYIIVPSTFEENKGDPCIRVPESEKKADYQAVDDEIENLEBPISDIDDDGVRRLPAQ 540
Qy 541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
Db 541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
Qy 601 WTQIOKQKIRYREIDVDGSGTMSYEMKALEAGFKNPCQLHOVYVARFADQIILDFD 660
Db 601 WTQIOKQKIRYREIDVDGSGTMSYEMKALEAGFKNPCQLHOVYVARFADQIILDFD 660
Qy 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDISWLCFSVL 700
Db 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDISWLCFSVL 700

```

RESULT 2

```

Qy 09YIC1 PRELIMINARY; PRT; 705 AA.
AC 09YIC1;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
Db Quail calpain.
GN OCLAL-1.
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Maeda Y.;
RT "cDNA sequence of quail skeletal muscle calpain."
DR EMBL; AB011080; BAA74564.1; -.
DR HSSP; P04574; IALV.
DR MEROPS; C02.003; -.

```

```

DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR000168; Shprot_acsite.
DR Pfam; PF001067; Calpain_I1; 1.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00648; peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_I1; 1.
DR SMART; SM00230; Cyapc_I.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR SEQUENCE 705 AA; 80188 MW; 5C70634EB3AA93CD CRC64;

```

Query Match 69.2%; Score 2568.5; DB 13; Length 705;
Best Local Similarity 65.9%; Pred. No. 1e-186; Matches 121; Indels 1; Gaps 1;
Matches 460; Conservative 116; Mismatches 121; Indels 1; Gaps 1;

```

Qy 3 GIAAKLAKDREAEGLSHERAIKYNODYEALRNECLAEAGTLFODPSFPAISALGFKE 62
Db 6 GIAARLORDLRABEGVEHNNAYKYNODYEALKQACIESGALFRDQFAGFTALGFKE 65
Qy 63 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 122
Db 66 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 125
Qy 123 RAVPLNOSFOENAGIFHFQFQWQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 182
Db 126 RAVPHGGSFOEDYAGIFHFQFQWQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 185
Qy 183 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIQXALQKSLGCGSI 242
Db 186 KAKAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIQXALQKSLGCGSI 245
Qy 243 DITSADSEALITFOKLVKHAIVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDNCP 302
Db 246 DITSAPMEAVITFOKLVKHAIVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDNCP 305
Qy 303 SMTVIDEERERLTRHEDGEFMSFSDFLRHYSRLICNLTPDILTSDTYKKMKLTMD 362
Db 306 EMDVIDEERERLTRHEDGEFMSFSDFLRHYSRLICNLTPDILTSDTYKKMKLTMD 365
Qy 363 GNNRGSTAGGCNRYNPTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRRRQRKM 422
Db 366 GNNRGSTAGGCNRYNPTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRRRQRKM 425
Qy 423 GEDMHTIGFYIYVEPEELSGQTNILSKNPFLLNRAERSDFTINREVLNRPKLPG 482
Db 426 GEDMHTIGFYIYVEPEELSGQTNILSKNPFLLNRAERSDFTINREVLNRPKLPG 485
Qy 483 IIVPSTFEENKGDPCIRVPESEKKADYQAVDDEIENLEBPISDIDDDGVRRLPAQ 541
Db 486 IIVPSTFEENKGDPCIRVPESEKKADYQAVDDEIENLEBPISDIDDDGVRRLPAQ 545
Qy 542 AGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 601
Db 546 AGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 605
Qy 602 TKIQKQKIRYREIDVDGSGTMSYEMKALEAGFKNPCQLHOVYVARFADQIILDFD 661
Db 606 NKIRSWLTIROYDLDSGTSYEMKALEAGFKNPCQLHOVYVARFADQIILDFD 665
Qy 662 FVCLVRLFTLFIKIQOLDPENTGTIELDISWLCFSV 699
Db 666 FVCLVRLFTLFIKIQOLDPENTGTIELDISWLCFSV 703

```

RESULT 3

```

Qy 08UW96 PRELIMINARY; PRT; 724 AA.
AC 08UW96;
DT 01-MAR-2002 (Tremblrel. 20, Created)

```


DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE mu/m-calpain large subunit.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyoshi K., Sorimachi H., Tomioka S., Ishiura S., Suzuki K.;
 RT "Xenopus has a calpain most similar to mu/m-type chicken calpain.",
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061521; BAB83262.1; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000169; SHprot acsite.
 DR Pfam; PF01067; Calpain_IIT; 1.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_IIT; 1.
 DR SMART; SM00054; Eph; 2.
 DR SMART; SM00054; Eph; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 SQ SEQUENCE 724 AA; 81770 MW; CA5CB090AE7A4288 CRC64;

Query Match 66.8%; Score 2480; DB 13; Length 724;
 Best Local Similarity 63.4%; Pred. No. 5.9e-180;
 Matches 455; Conservative 115; Mismatches 126; Indels 22; Gaps 3;

QY 3 GIAAKLAKDREAAGLGSHERAIKYLNQDYALRNECLEAGTLFODSPPAIPALGFKE 62
 DB 6 GIASKLIKIDLKSGVSHOCAYKQNDYSLKQCVESGILFEDINFAIPISILGFKE 65
 QY 63 LGFYSSKTRGMWRKPRTEICADPOFIIGATRTDICOAGLDCWLLAALASLTINEILA 122
 DB 66 LGFSSSKTRGVQWPRSDIYDDPQFIIGATRTDICOAGLDCWLLAALASLTINEIDLH 125
 QY 123 RVPLNOSPOENYAGTFHPOFQWGEVWVVDRLPTQXGELLFPHSAGSEFWSLTE 182
 DB 126 RVVPHGSGFQEDYAGTFHPOFQWGEVWVVDRLPTQXGELLFPHSAGSEFWSLTE 185
 QY 183 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKRPENLFFKIILOKLSLGCST 242
 DB 186 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKRPENLFFKIILOKLSLGCST 245
 QY 243 DITSADSEAITTQKLVKGHAISVTGAEEVSNGLQKLRIRNPGWEVMTGRMNDNC 302
 DB 246 DITSADSEAITTQKLVKGHAISVTGAEEVSNGLQKLRIRNPGWEVMTGRMNDNC 305
 QY 303 SMTTIDEEERELTRRHEDEEFWMSFDFLRHYSRLICNLTPPTLTSDTYKKKLTMD 362
 DB 306 EMEVDSSEEDRLRKMEDEEFWMSFDFLRHYSRLICNLTPPTLTSDTYKKKLTMD 365
 QY 363 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDDESGGCTFLVGLIOKRRROR 422
 DB 366 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDDESGGCTFLVGLIOKRRROR 425
 QY 423 GEDMTTIGFQIYVEBEELSGQTNHLSKNPFLTRNARSDDTFLNREVLNRFPLPGEY 482
 DB 426 GEDMTTIGFQIYVEBEELSGQTNHLSKNPFLTRNARSDDTFLNREVLNRFPLPGEY 485
 QY 483 IIVPSTPEPKKDDFCIRVSEKKADYQAVDDLEAUL-EEFDSSEDDIDGVRRLAQL 541
 DB 486 IIVPSTPEPKKDDFCIRVSEKKADYQAVDDLEAUL-EEFDSSEDDIDGVRRLAQL 545
 QY 542 AGEDAIISAFELQTLIRVLAKRODIDSGSIFETCKIMVMDLSDSGKLGKEFYILM 601
 DB 546 AGEDAIISAFELQTLIRVLAKRODIDSGSIFETCKIMVMDLSDSGKLGKEFYILM 605
 QY 602 TKIQ-----KYQIYREIDVDRSGTWNSEYMRKALEAGFKMPCQ 641

DB 606 NKEFAMVALLPACMGPRVGINTPSLQTVFEREDLKSIGTSSYEIRALNESSGYKLNNK 665
 QY 642 LHOIVARFADDDLIIDFNFVACLVRLTELFPIFQLODPENGTIELDLSMLCFSY 699
 DB 666 LIQVAVRYADSDMGIDFDFVCCLVKLEAMFPRFALD-EGDGTAEMLNGEMLTMTM 722

RESULT 4
 09GLG2 PRELIMINARY; PRT; 714 AA.

AC 09GLG2; DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Calpain 1.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NC NCB1_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakajima T., Fukiage C., Azuma M., Shearer T.R.;
 RT "Calpain isoforms in the eye of monkey."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF284440; AAG22770.1; -.
 DR HSSP; P04574; 1ALV.
 DR MEROPS; C02.001; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000169; SHprot acsite.
 DR Pfam; PF01067; Calpain_IIT; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_IIT; 1.
 DR SMART; SM00230; CysPc; 1.
 DR SMART; SM00054; Eph; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 SQ SEQUENCE 714 AA; 81849 MW; 834690C14DE7AD8 CRC64;

Query Match 66.4%; Score 2464.5; DB 6; Length 714;
 Best Local Similarity 63.2%; Pred. No. 8.7e-179;
 Matches 440; Conservative 124; Mismatches 129; Indels 3; Gaps 2;

QY 3 GIAAKLAKDREAAGLGSHERAIKYLNQDYALRNECLEAGTLFODSPPAIPALGFKE 62
 DB 13 GVSAGVQKQAKELGGRHENAIXYLGQDYEQRLARCLQSGTLFRDEAPFPVQSLGFKD 72
 QY 63 LGFYSSKTRGMWRKPRTEICADPOFIIGATRTDICOAGLDCWLLAALASLTINEILA 122
 DB 73 LGFSSSKTRGVQWPRSDIYDDPQFIIGATRTDICOAGLDCWLLAALASLTINEIDLH 125
 QY 123 RVPLNOSPOENYAGTFHPOFQWGEVWVVDRLPTQXGELLFPHSAGSEFWSLTE 182
 DB 133 RVVPHGSGFQEDYAGTFHPOFQWGEVWVVDRLPTQXGELLFPHSAGSEFWSLTE 192
 QY 183 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKRPENLFFKIILOKLSLGCST 242
 DB 193 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKRPENLFFKIILOKLSLGCST 252
 QY 243 DITSADSEAITTQKLVKGHAISVTGAEEVSNGLQKLRIRNPGWEVMTGRMNDNC 302
 DB 253 DITSADSEAITTQKLVKGHAISVTGAEEVSNGLQKLRIRNPGWEVMTGRMNDNC 312
 QY 303 SMTTIDEEERELTRRHEDEEFWMSFDFLRHYSRLICNLTPPTLTSDTYKKKLTMD 362
 DB 313 SMTTIDEEERELTRRHEDEEFWMSFDFLRHYSRLICNLTPPTLTSDTYKKKLTMD 372
 QY 363 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDDE--DGSGCTFLVGLIOKRRROR 420


```

Db 493 PGEYVVVSTEEPNKGDGVLFVRFSEKSGAGTQELDDQVQANLPDEQVISEEDIDENFKSL 552
Qy 538 FAOLAGEAEISAFELQTLIRVLAKRODIDKSDGSIETCKIMVMDLSDSGKLGKEF 597
Db 553 FROLAGEMEISVVELRTILNLTISKHDLRTNGFSLSCSRMVMDLRDNGKLGVEFNI 612
Qy 598 YLWTKIOYOKIYREIDVDRSGTMMNSYEMKALAEAGFKMPQOLHQVTVAFADQLIIF 657
Db 613 NILWNRINNYLSTIFKFPDLKSGMSAYEMEMALIEFAPFKLNKKLYELIITRYSRPLAV 672
Qy 658 DFDNVRCLVRLTETFKIFKQOLDPENTGTIELDLISWL 695
Db 673 DFDNFCCLVRLTETFWRFKFLDLDGVVTFDLFKWL 710

```

RESULT 8

```

ID 08C2J1 PRELIMINARY; PRT; 713 AA.
AC 08C2J1;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ND; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK088547; BAC40416.1;
SQ SEQUENCE 713 AA; 82152 MW; 3E1FC84D5802B864 CRC64;

```

```

Query Match 65.3%; Score 2426; DB 11; Length 713;
Best local Similarity 62.6%; Pred. No. 7.4e-176;
Matches 436; Conservative 126; Mismatches 130; Indels 4; Gaps 3;

```

```

Qy 3 GIAAKLADREAAEGLGSHERAIKYLNQDYALRNECLEAGTLFQDPSPAPISALGKFE 62
Db 13 GVSAGVOVKRDKELGLGHENAIKTLGQDYETLRACISGVLFDDEAFPPVSHSLGRKE 72
Qy 63 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAISLTNEBIIA 122
Db 73 LGPHSSKTYGIKMKRPTEIMSNPQFIYDGAIRTDICQALGDCWLLAIAISLTNETILH 132
Qy 123 RVPLNGSFQENYAGIFHFQWQYGEWVYVVDRLPTQDGLLFVHSAESSEFSALLE 182
Db 133 RVPPGGSFQDYGAGIFHFQWQYGEWVYVVDRLPTQDGLLFVHSAQNEFSALLE 192
Qy 183 KAYAKINGCYEALSGATTEGFEFTGIAEWYELKKPPNLFKIIQALQKSLGCSI 242
Db 193 KAYAKINGSYALSGCCSEAFEDFTGVTWYDQKAPSDUIYQIILKALERSGLGCSI 252
Qy 243 DITSAADEALTPQKLVKGAHAYVTGAEVESNGSLQKILIRNPWGEVETGRWINDCP 302
Db 253 NISDIRLEAITPQKLVKGAHAYVTGAKOVITYQGGORVVLIMRNPNWGEVEKGPMSDSY 312
Qy 303 SMNTIDPEERERLRKHEDGFEWMSFSPLHYSRLCINTPTDLSDTYKKMKLTAMD 362
Db 313 EWNVNDYPERQLRKWEDGFEWMSFRDPIREFTKLEICNLTPLAKSRITLRNNTTFYE 372
Qy 363 GNMRRGSTAGGCRNYPNTFMNPOYLKLEBED--EDEDESGSGTFLVGLIKHRROR 420
Db 373 GTMRGSGTAGGCRNYPNTFMNPOYLKLEBED--EDEDESGSGTFLVGLIKHRROR 432
Qy 421 KMGEDMTIGGIYEVPEELSGQTNILSKNPFILNRARSDTFINLRVNLRFKLPFG 480

```

```

Db 433 RFGHDMETIGFAYQVPRELAGQP-VHLKQDFLANASRAQSEHFILNLRVNRIRLRPG 491
Qy 481 EYLIVSTPEPNKGDGDCIVFSEKADQVAVDEIEMNL-EEEDISEDDIDGVRLLFA 539
Db 492 EYLIVSTPEPNKGDGDLVLFSEKAGTQELDDQVQANLPDEKIVSEEDIDENFKLFS 551
Qy 540 QLAGEDAEISAFELQTLIRVLAKRODIDKSDGSIETCKIMVMDLSDSGKLGKEFYI 599
Db 552 KLAGEDAEISVVELRTILNLTISKHDLRTNGFSLSCSRMVMDLRDNGKLGVEFNI 611
Qy 600 LWTIKIYOKIYREIDVDRSGTMMNSYEMKALAEAGFKMPQOLHQVTVAFADQLIIF 659
Db 612 LWNIRINNYLSTIFKFPDLKSGMSAYEMEMALIEFAPFKLNKKLYELIITRYSRPLAV 671
Qy 660 DNFVRCCLVRLTETFKIFKQOLDPENTGTIELDLISWL 695
Db 672 DNFVCCCLVRLTETFWRFKFLDLDGVVTFDLFKWL 707

```

RESULT 9

```

ID 0918G2 PRELIMINARY; PRT; 704 AA.
AC 0918G2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain 1 (EC 3.4.22.17) (Danio rerio).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee K.C., Evans C.W.;
RT "Zebrafish calpain 1 (camp1): identification, expression and
RL phylogenetic implications.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282675; AAF82808.1;
DR HSSP; P04574; ALV.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR00169; SH3prot acsite.
DR Pfam; PF01067; Calpain III; 1.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00648; peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain III; 1.
DR SMART; SM00230; Cyapc; 1.
DR SMART; SM00054; Efp; 2.
DR PROSITE; PS00018; EF_HAND_2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR KEGG; K01039; THIOL_PROTEASE_CYS; 1.
SQ SEQUENCE 704 AA; 80090 MW; E5FBC08C18A79A8 CRC64;

```

```

Query Match 64.1%; Score 2381; DB 13; Length 704;
Best local Similarity 62.5%; Pred. No. 1.9e-172;
Matches 436; Conservative 105; Mismatches 153; Indels 4; Gaps 2;

```

```

Qy 3 GIAAKLADREAAEGLGSHERAIKYLNQDYALRNECLEAGTLFQDPSPAPISALGKFE 62
Db 8 GMAARLRSCMDRAGAGQNNNAKPIFGQDYETLRAGSQSRRRPFEDPMFLAASSSLGFNE 67
Qy 63 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAISLTNEBIIA 122
Db 68 LGPHSSKTYGIKMKRPTEIMSNPQFIYDGAIRTDICQALGDCWLLAIAISLTNETILH 127
Qy 123 RVPLNGSFQENYAGIFHFQWQYGEWVYVVDRLPTQDGLLFVHSAESSEFSALLE 182
Db 128 RVCRTGQDFSRVAGIFHFQWQYGEWVYVVDRLPTQDGLLFVHSAESSEFSALLE 187
Qy 183 KAYAKINGCYEALSGATTEGFEFTGIAEWYELKKPPNLFKIIQALQKSLGCSI 242

```

```

Db 188 KAAKANGCEALSGSGSTGEGFDTGVTGEMWELKAPADLPISIGRAIERSLGCGSI 247
Qy 243 DITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRANDNCP 302
Db 248 DITSKEDMEANTFKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRANDNCP 307
Qy 303 SWNTIDPEERERLRRHEDGFMWFSDFLRHYSRLKICNLTPDPLTSDTYKKCKLTKMD 362
Db 308 EMDNDRSVGRGLONRSEDEGFWMWFSDFLRHYSRLKICNLTPDPLTSDTYKKCKLTKMD 367
Qy 363 GNNRRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRRORRM 422
Db 368 GEMRRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRRORRM 424
Qy 423 GEMMTIGFIVPEELSGOTNHLKSNFELTNRAERSDTFLNREVLNPFLLPGEY 482
Db 425 GOMETIGFIVPEELSGOTNHLKSNFELTNRAERSDTFLNREVLNPFLLPGEY 484
Qy 483 ILVSTPEPNKDDPCIRVFSEKKADYQAVDEIEANTL-EEFDISEDDIDDGVARLPAL 541
Db 485 ILVSTPEPNKDDPCIRVFSEKKADYQAVDEIEANTL-EEFDISEDDIDDGVARLPAL 544
Qy 542 AGEDAEISAELOTLIRVLAKRODYSDFSIETCKINVDMLDSDSGGLKEPITLM 601
Db 545 AGADMEISVTELOTLIRVLAKRODYSDFSIETCKINVDMLDSDSGGLKEPITLM 604
Qy 602 TKLOKOKIYREIDVRSFGTNSYEMKALBAGFKPKCOLHOIVARFADPOLIIDFN 661
Db 605 EKIIRYQIYREIDVRSFGTNSYEMKALBAGFKPKCOLHOIVARFADPOLIIDFN 664
Qy 662 FVRCLVLETLFKIYREIDVRSFGTNSYEMKALBAGFKPKCOLHOIVARFADPOLIIDFN 699
Db 665 FVSCIVLETLFKIYREIDVRSFGTNSYEMKALBAGFKPKCOLHOIVARFADPOLIIDFN 702

```

RESULT 10

Q64698 PRELIMINARY; PRT; 703 AA.

```

AC 064698;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain, large (Catalytic) subunit (EC 3.4.22.17) (Calcium-activated
DE neutral proteinase) (CAMP) (Stomach-specific calcium-activated neutral
DE protease large subunit) (NCCL2).
GN CUS4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH;
RX MEDLINE=93374936; PubMed=7690035;
RA Sorimachi H., Ishiura S., Suzuki K.;
RT "A novel tissue-specific calpain species expressed predominantly in
RT the stomach comprises two alternative splicing products with and
RT without Ca(2+)-binding domain";
RL J. Biol. Chem. 268:19476-19482(1993).
CC - FUNCTION: CALPAINS ARE CALCIUM-ACTIVATED NON-LYSOSOMAL THIO-
CC PROTEASES.
CC - SUBUNIT: HETERODIMER OF A LARGE (CATALYTIC) AND A SMALL
CC (REGULATORY) SUBUNIT.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; NCCL2 (SHOWN HERE) AND NCCL2', ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE STOMACH. LOW IN
CC INTESTINE AND HEART.
CC - MISCELLANEOUS: IN RAT THERE SEEMS TO BE 2 TYPES OF CALPAIN-
CC UBQUITOUS FORMS, CALPAIN I (MICRO-MOLDS CA++ REQUIRING) AND
CC CALPAIN II (MILI-MOLE CA++ REQUIRING), AND TISSUE SPECIFIC FORMS,
CC CALPAIN P94 AND NCCL2. THE SMALL UNIT IS COMMON TO ALL FORMS.

```

```

CC - MISCELLANEOUS: THIS PROTEIN SEEMS TO BIND TWO MOLES OF CALCIUM.
CC - SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2; ALSO KNOWN AS THE
CC CALPAIN FAMILY OF THIOL PROTEASES.
DR EMBL; D14478; BA03369.1; -
DR EMBL; D14479; BA03370.1; -
DR EMBL; D14480; BA03371.1; -
DR HSSP; P04574; 1ALV.
DR MEROPS; C02.007; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Stpoc_acetate.
DR InterPro; IPR000169; Stpoc_acetate.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; CyBpc; 1.
DR SMART; SM00054; Efn; 2.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR Hydrolase; Thiol protease; Calcium-binding; Multigene family;
KW Alternative splicing.
FT DOMAIN 1 73
FT 74 332
FT 333 532
FT DOMAIN 533 703
FT ACT_SITE 105 105
FT ACT_SITE 262 262
FT ACT_SITE 286 286
FT CA_BIND 588 599
FT CA_BIND 618 629
FT DOMAIN 653 664
FT DOMAIN 683 694
FT VARSPIC 103 103
FT VARSPIC 380 380
FT VARSPIC 381 703
SQ SEQUENCE 703 AA; 79555 MW; C0688B055FC0DEC CRC64;

Query Match 63.5%; Score 2356.5; DB 11; Length 703;
Best Local Similarity 60.9%; Pred. No. 1.4e-170; Indels 7; Gaps 3;
Matches 427; Conservative 127; Mismatches 140;

Qy 1 MAGIAKLAKDRRAAAGLSHERAIXYLNQDYALNNECEAGTLQDPSFPAIPALGF 60
Db 1 MALAAGVSKQRAVAGLSGNQNAVXYLGDPETLAKQCLNSVLRKDPFRPCBALGF 60
Qy 61 KELGYSKTRGRMRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTINEI 120
Db 61 KDLGSPSPDTGIVMKRPTELCPNPQIVGATRTDIPQGLDGCWLLAIAISLTINEI 120
Qy 121 LARVPLNGSPQNNVAGIFHFQFWQGEWVVDRLPTKDELLFVHABSESEFMSAL 180
Db 121 LYRVLPRDPSFQKDVAGIFHFQFWQGEWVVDRLPTKDELLFVHABSESEFMSAL 180
Qy 181 LEKAYAKINGCYALSGATTEGFEDFTGGAIEWELKRPPLFKTIQKALQKSLGC 240
Db 181 LEKAYAKINGCYALSGATTEGFEDFTGGAIEWELKRPPLFKTIQKALQKSLGC 240
Qy 241 SIDITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRANDN 300
Db 241 SIDITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRANDN 300
Qy 301 CPBWNTIDPEERERLRRHEDGFMWFSDFLRHYSRLKICNLTPDPLTSDTYKKCKLTK 360
Db 301 CPBWNTIDPEERERLRRHEDGFMWFSDFLRHYSRLKICNLTPDPLTSDTYKKCKLTK 360
Qy 361 MDGNRRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRR 418
Db 361 MDGNRRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRR 418
Qy 419 QRMKGEDMTIGFIVPEELSGOTNHLKSNFELTNRAERSDTFLNREVLNPFLLP 478
Db 419 QRMKGEDMTIGFIVPEELSGOTNHLKSNFELTNRAERSDTFLNREVLNPFLLP 478

```

Db 421 QKRIGQMLSDIGAVVOYIPKELSHSTDAHLGRDFELGRQPSSTCYMNLREVSRYRLP 480
Qy 479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLEEF--FDISEDDIDGVR 535
Db 481 PGOYLVPSTPEPRKDDFCIRVFSEKKALIEIGDYVSGPHPEPRMDME--DEHYR 538
Qy 536 RLFAQLAGEDEAIEAFELQTLIRRLVLAARODIKSDGFSIETCKIMVMDLSDSGSKLGLK 595
Db 539 SLFEFVGKDEISEANQLKRYLNEVLKRTDMKFDGFMINTCREMISLSDSGTSGICPM 598
Qy 596 EFTYLMKIKYOKYKRYREIDVDKSGTMSYEMRKALBEAGKRCQHLQHVAVFADDOL 655
Db 599 EFTYLMKIKRYLYLIPQEMDNHVGTLFAHMRTRALKKAGFTLNNQVQTLAMRYACSKL 658
Qy 656 IIDFDFVRCIVRLETLFKIFKQDPENTGTIELDLISMLC 696
Db 659 GVDINGFVACMIRLETLFKLFRLLDKONGIYQSLAEMLC 699

RESULT 11

Q8BPV9 PRELIMINARY; PRT; 530 AA.
AC Q8BPV9; 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MBLINB=22354683; Pubmed=12466851;
RA the FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK052141; BAC34855.1; -
DR SEQUENCE 530 AA; 5929 MM; 7BD6DCDCC1482BC92 CRC64;
SQ
Query Match 62.9%; Score 2335; DB 11; Length 530;
Best Local Similarity 71.0%; Pred. No. 4e-169;
Matches 451; Conservative 24; Mismatches 28; Indels 132; Gaps 2;

Qy 1 MAGIAAKIADKREAAEGISHERAIKYLINDYALRNECLAGTLFQDPSPPAISALGF 60
Db 1 MAGIAIKIADKREAAEGISHERAIKYLINDYETLRNECLBAGALFQDPSPPALPSSIGY 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDIOGALDGCWILAAIASLTINEI 120
Db 61 KELGPYSSKTRGIMWKRPTEICADPOFIIGATRTDIOGALDGCWILAAIASLTINEI 120
Qy 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 181 LEKAYAKINGCYEALSGGATTEGFEFTGIAEWYELKPPPNLFEKIIQKALOKSGSLGC 240
Db 181 LEKAYAKINGCYEALSGGATTEGFEFTGIAEWYELKPPPNLFEKIIQKALOKSGSLGC 240
Qy 241 SIDITSAADSRAITFOKLVKHAISVYGAEEVESGSIQKILIRNPMGEVEMTKRWNND 300
Db 241 SIDITSAADSRAITFOKLVKHAISVYGAEEVESGSIQKILIRNPMGEVEMTKRWNND 300
Qy 301 CPSNNTIDPEERELTRHEDGEPFMSFSLRAHSRLICNLPPDITSDTYKKWKLT 360
Db 301 CPSNNTIDPEERELTRHEDGEPFMSFSLRAHSRLICNLPPDITSDTYKKWKLT 360
Qy 361 MDGNMRBSSTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420

Db 359 ----- 358
Qy 421 KMGEDMHTIGFIYEVEBELSGQTNHLSKNFELTNRAERSDPTINLREVLNRFKLP 480
Db 359 ----- 358
Qy 481 EYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLEEFDEIDIDGVRRLFAQ 540
Db 359 -----TFPRHKGDFCIRVFSEKKADYQAVDEIEANLEEFIDANEEDIDGFRRLFVQ 411
Qy 541 LAGEDAIEISAFELQTLIRRLVLAARODIKSDGFSIETCKIMVMDLSDSGSKLGLKXEPYL 600
Db 412 LAEBDEAIEISAFELQTLIRRLVLAARODIKSDGFSIETCKIMVMDLSDSGSKLGLKXEPYL 471
Qy 601 WTKIYOKYKRYREIDVDKSGTMSYEMRKALBEAG 635
Db 472 WTKIYOKYK--NLPGRCQGVNHHETLRDAESTG 503

RESULT 12

Q91VA3 PRELIMINARY; PRT; 703 AA.
AC Q91VA3; 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Stomach specific calpain nCL-2 (Stomach-specific calpain).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hata S., Sorimachi H., Lee H., Kawahara H., Maeda T., Suzuki K.;
RT "Both conserved and unique gene structure of stomach-specific calpains
RT revealed a process of calpain gene evolution."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Hata S., Nishi K., Kawamoto T., Lee H., Kawahara H., Maeda T.,
RA Shintani Y., Sorimachi H., Suzuki K.;
RT "Both the conserved and unique gene structure of stomach-specific
RT calpains reveal processes of calpain gene evolution."
RL J. Mol. Evol. 0:0-0(2001).
DR EMBL; AB050202; BAB70480.1; -
DR EMBL; AB061518; BAB55000.1; -
DR MEROPS; C02.007; -
DR MGD; MGI:2181366; Capn8.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR001659; SHprot_acsite.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00648; peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00220; Cyapc; 1.
DR SMART; SM00054; Efn; 3.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
SQ SEQUENCE 703 AA; 79334 MM; 50F2P9CCF2192FB6 CRC64;

Query Match 62.6%; Score 2322.5; DB 11; Length 703;
Best Local Similarity 60.2%; Pred. No. 5.5e-168;
Matches 421; Conservative 131; Mismatches 144; Indels 3; Gaps 2;

Qy 1 MAGIAAKIADKREAAEGISHERAIKYLINDYALRNECLAGTLFQDPSPPAISALGF 60
Db 1 MAGIAIKIADKREAAEGISHERAIKYLINDYETLRNECLBAGALFQDPSPPALPSSIGY 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDIOGALDGCWILAAIASLTINEI 120
Db 61 KELGPYSSKTRGIMWKRPTEICADPOFIIGATRTDIOGALDGCWILAAIASLTINEI 120
Qy 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQFOWYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 181 LEKAYAKINGCYEALSGGATTEGFEFTGIAEWYELKPPPNLFEKIIQKALOKSGSLGC 240
Db 181 LEKAYAKINGCYEALSGGATTEGFEFTGIAEWYELKPPPNLFEKIIQKALOKSGSLGC 240
Qy 241 SIDITSAADSRAITFOKLVKHAISVYGAEEVESGSIQKILIRNPMGEVEMTKRWNND 300
Db 241 SIDITSAADSRAITFOKLVKHAISVYGAEEVESGSIQKILIRNPMGEVEMTKRWNND 300
Qy 301 CPSNNTIDPEERELTRHEDGEPFMSFSLRAHSRLICNLPPDITSDTYKKWKLT 360
Db 301 CPSNNTIDPEERELTRHEDGEPFMSFSLRAHSRLICNLPPDITSDTYKKWKLT 360
Qy 361 MDGNMRBSSTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420
Db 361 MDGNMRBSSTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420

QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLEKKPPNLFKIQALOKSILGC 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLEKKPPNLFKIQALOKSILGC 240
 QY 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 DB 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 QY 301 CPSEWNTIDPEERELTRRHDEGEFMSFSDFLRHYSRLICNLTPTDITSDTYKKWLTJK 360
 DB 301 CPSEWNTIDPEERELTRRHDEGEFMSFSDFLRHYSRLICNLTPTDITSDTYKKWLTJK 360
 QY 361 MDGWMRSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--ESGCTFLVGLIQKRRR 418
 DB 361 MDGWMRSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--ESGCTFLVGLIQKRRR 418
 QY 419 QKRGEDMTIGFQIYVEPEELSGQTNILSKNFFLTNRARERSDTFNLREVLRPKLP 478
 DB 419 QKRGEDMTIGFQIYVEPEELSGQTNILSKNFFLTNRARERSDTFNLREVLRPKLP 478
 QY 421 QRRGQMLSTIGAVYQIPKLENTDEHILGRDFQGRQSTSCSTYMLREVSSRVQLP 480
 DB 421 QRRGQMLSTIGAVYQIPKLENTDEHILGRDFQGRQSTSCSTYMLREVSSRVQLP 480
 QY 479 PGEYILVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 537
 DB 479 PGEYILVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 537
 QY 538 PAOLAGDAISAFELQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 597
 DB 538 PAOLAGDAISAFELQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 597
 QY 541 SEEDADDSSEISAHQLRVINGLSKRTDMKDFGNTCREMISLIDGQGTSLRVEF 600
 DB 541 SEEDADDSSEISAHQLRVINGLSKRTDMKDFGNTCREMISLIDGQGTSLRVEF 600
 QY 598 YILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 657
 DB 598 YILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 657
 QY 601 KTLMLKICXLEIYQENHDSBAGTIIDAHKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 DB 601 KTLMLKICXLEIYQENHDSBAGTIIDAHKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 QY 658 FDNFVRCVLRLETLFKIQOLDPENTGTIELDLISWLC 696
 DB 658 FDNFVRCVLRLETLFKIQOLDPENTGTIELDLISWLC 696
 QY 661 DFDGFVACMIRLETLFKRLDQNGIVQLSLAEWLC 699
 DB 661 DFDGFVACMIRLETLFKRLDQNGIVQLSLAEWLC 699

RESULT 13

Q918T0 PRELIMINARY; PRT; 702 AA.
 AC Q918T0
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Calpain.
 GN CL-2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCBITaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cao Y., Zhao H., Grunz H.
 RT "A novel Xenopus gene homologous to rat calpain."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF212199; AAF63194.2; -
 DR HSSP; P04574; 1ALV.
 DR MEROPS; C02.007; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHProl_acsite.
 DR Pfam; PF01067; Calpain_IIT; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN_IIT; 1.
 DR SMART; SM00720; calpain_IIT; 1.
 DR SMART; SM00230; Cyapc; 1.

DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR SEQUENCE 702 AA; 7946 MW; 6D6A3B1D276DC51 CRC64;
 Query Match 62.2%; Score 2309; DB 13; Length 702;
 Best Local Similarity 60.7%; Pred. No. 5.8e-167;
 Matches 426; Conservative 119; Mismatches 155; Indels 2; Gaps 1;

QY 1 MAGIAKTLADREAEGLSGHERAIKYLNDYALNRECLEAGTLFQDPFPAPLPSALGF 60
 DB 1 MMSAANIAADRLADGCGTKRKPETKLDDEFKLAQCLASGALYDEFPACPSALGY 60
 QY 61 KELGPYSSKTRGMRKRPTEICADPOITIGATRTDICGALDCWLLAIALVLTNEE 120
 DB 61 NEILRPGSYKTSGLVWKRPTEICPNPQFIVDGAATRGDIRQALDCWLLAIALVLTNEE 120
 QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLEKKPPNLFKIQALOKSILGC 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLEKKPPNLFKIQALOKSILGC 240
 QY 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 DB 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 QY 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 DB 241 SIDTSAADSEAITFOKLVKHAAYVTGAEEVESNGSLQKILIRINPMGEVETGWRWN 300
 QY 301 CPSEWNTIDPEERELTRRHDEGEFMSFSDFLRHYSRLICNLTPTDITSDTYKKWLTJK 360
 DB 301 CPSEWNTIDPEERELTRRHDEGEFMSFSDFLRHYSRLICNLTPTDITSDTYKKWLTJK 360
 QY 301 APEMNIVDPYKAVLNRKSEDEGFMAFSDFLREYSRLICNLTPTDITSDTYKKWLTJK 360
 DB 301 APEMNIVDPYKAVLNRKSEDEGFMAFSDFLREYSRLICNLTPTDITSDTYKKWLTJK 360
 QY 361 MDGWMRSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--DSEGCTFLVGLIQKRRR 418
 DB 361 MDGWMRSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--DSEGCTFLVGLIQKRRR 418
 QY 361 YTSMAKSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--DSEGCTFLVGLIQKRRR 420
 DB 361 YTSMAKSGTAGGCRNYPNTFMNPOVLKLEDEDEDEB--DSEGCTFLVGLIQKRRR 420
 QY 419 QKRGEDMTIGFQIYVEPEELSGQTNILSKNFFLTNRARERSDTFNLREVLRPKLP 478
 DB 419 QKRGEDMTIGFQIYVEPEELSGQTNILSKNFFLTNRARERSDTFNLREVLRPKLP 478
 QY 421 KKKMGEDLSIGSLFKIPQLODHTDAHILGRDFQKTPAASDFTYINVRVSSNRHLP 480
 DB 421 KKKMGEDLSIGSLFKIPQLODHTDAHILGRDFQKTPAASDFTYINVRVSSNRHLP 480
 QY 479 PGEYILVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 538
 DB 479 PGEYILVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 538
 QY 481 VGDYLVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 540
 DB 481 VGDYLVSTFEPNKGDFCIRVSEKKADYQAVDEIEANLEB-PDISDDIDGVRRL 540
 QY 539 AOLAGDAISAFELQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 598
 DB 539 AOLAGDAISAFELQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 598
 QY 541 DKLAGDEEVDALQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 600
 DB 541 DKLAGDEEVDALQTLIRVLAQRQDLSQSFSTETCKIMVMDSDSGKLGKEF 600
 QY 599 ILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 658
 DB 599 ILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 658
 QY 601 ILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 660
 DB 601 ILWTKIOKQIYREIDVRSQTMNSYEMKALEAGFKMPQOLHOVIVARFADDOIT 660
 QY 659 FDNFVRCVLRLETLFKIQOLDPENTGTIELDLISWLC 700
 DB 659 FDNFVRCVLRLETLFKIQOLDPENTGTIELDLISWLC 700
 QY 661 FDFGFIACMIRLETLFKRLDQNGIVQLSLAEWLC 702
 DB 661 FDFGFIACMIRLETLFKRLDQNGIVQLSLAEWLC 702

RESULT 14

Q9NOM7 PRELIMINARY; PRT; 650 AA.
 AC Q9NOM7
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Microtalar calcium-activated neutral protease I isoform B.
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NC NCBITaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21087425; PubMed=11219468;
 RA Smith T.P.L., Simmen F.A., Zhao G., Vallet J.L.;

"Nucleotide sequences of two isoforms of porcine micromolar calcium-RT activated neutral protease 1 cDNA."
RT J. Anim. Sci. 79:552-553 (2001).
DR EMBL: AF263609; AAF73443.1; -
DR HSSP: P04574; IALV.
DR MEROPS: C02.001; -
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001300; Protease_C2.
DR InterPro: IPR000169; SHprot_acsite.
DR Pfam: PF01067; Calpain_III.1.
DR Pfam: PF00036; ehand; 2.
DR Pfam: PF00648; peptidase_C2; 1.
DR PRINTS: PR00704; CALPAIN.
DR SMART: SM00720; calpain_III; 1.
DR SMART: SM00230; Cyapc; 1.
DR SMART: SM00054; Efh; 2.
DR PROSITE: PS00018; EF_HAND; 2.
DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
DR KEGG: K04486; THIOL_PROTEASE.
SQ SEQUENCE 650 AA; 74033 MW; 2EE9FF569D90FD7 CRC64;

Query Match 61.4%; Score 2279.5; DB 6; Length 650;
Best Local Similarity 64.5%; Pred. No. 9.1e-165;
Matches 411; Conservative 107; Mismatches 114; Indels 5; Gaps 3;

QY 3 GIAAKLAKDRAEAGLSHERAIKYLNDYALNNECLAGTLFQDPSPAPALGPK 62
DB 13 GVSQVQGLRAKELGLSHENAIKYLQDYEQLAHCHQSSSLRDEAFPPVPSLGKE 72
QY 63 LGYSSKTRGMRKRPETICADPOFIIIGATRTDICOALDCCWLLAIASTLTNEIIA 122
DB 73 LGPSSSKTYGVKMRPTELFSNPQFIYDGAATRTDICOALDCCWLLAIASTLTNDLIIH 132
QY 123 RVVLNDSFOENYAGIFHFQFQMGWEVVDRLPTKDELLFVHSAGSEFMSALE 182
DB 133 RVPHGQSFQNGYAGIFHFQFQMGWEVVDRLPTKDELLFVHSAGSEFMSALE 192
QY 183 KAVAKINGCYEALSGGATTEGFEDFTGIAEMYLKRPNNLKKIIOKALQKSLGCSI 242
DB 193 KAVAKINGCYEALSGGATTEGFEDFTGIAEMYLKRPNNLKKIIOKALQKSLGCSI 252
QY 243 DITSAADSEAITFOKLKVGHAIVTGAEEVSNGLQKLIIRNPMGEVETGRANDNC 302
DB 253 DISSVLDMEAIVTFOKLKVGHAIVTGAEEVSNGLQKLIIRNPMGEVETGRANDNC 312
QY 303 SMNITDPERERLTPRHEDGFWMSFSDLRHSLEICNLTPDLTSDTYKKMKLTAMD 362
DB 313 EWNQVDYQDQLRVRMEDGFWMSFSDLRHSLEICNLTPDLTSDTYKKMKLTAMD 372
QY 363 GNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEED---GEGCTFLVGLIOKRRRO 419
DB 373 GNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEED---GEGCTFLVGLIOKRRRO 432
QY 420 RKMGEDHTIGFGLIYEPPELISGQTNHLSKNFLLNTRARSDTFINLREVLNFKLP 479
DB 433 RRFQDMETIIFAVAEVPELVGP-VHLKEDFFLANASRASEQFINLREVSFRLP 491
QY 480 GEYLIVSTFEPNKGDCIFVSEKADYQAVDEIEANL-EEEDISEDIDDOVRLLF 538
DB 492 GEYLIVSTFEPNKGDCIFVSEKADYQAVDEIEANL-EEEDISEDIDDOVRLLF 551
QY 539 AQLAGEDAEISAFELQTLIRRLARODIKSDGSIETCKIMVMDLSDSGKLGKEFY 598
DB 552 RQLAGEDEISAFELQTLIRRLARODIKSDGSIETCKIMVMDLSDSGKLGKEFY 611
QY 599 ILMTKIQYQKIYREIVDNRSGTMSYEMKALAEAG 635
DB 612 ILMTKIRYVLSIFRKFDLSDGSMAYEMKALAEAG 648

RESULT 15
Q8R486 PRELIMINARY; PRT; 421 AA.

AC Q8R486;
DT 01-JUN-2002 (TREMURel. 21, Created)
DT 01-JUN-2002 (TREMURel. 21, Last sequence update)
DT 01-MAR-2003 (TREMURel. 23, Last annotation update)
DE M-calpain 80 kDa large subunit (Fragment).
GN CAPN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Elce J.S., Arthur J.S.C., Croall D.E., Dutt P.;
RT "Mouse m-calpain, partial genomic sequence."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF97625; AM19226.1; -
DR InterPro: IPR001300; Protease_C2.
DR Pfam: PF01067; Calpain_III; 1.
DR Pfam: PF00648; peptidase_C2; 1.
DR SMART: SM00720; calpain_III; 1.
DR SMART: SM00230; Cyapc; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 421 AA; 48334 MW; 271E450B7B95E633 CRC64;

Query Match 57.7%; Score 2142; DB 11; Length 421;
Best Local Similarity 93.3%; Pred. No. 1.4e-154;
Matches 393; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 143 FMQGEVWEVVDRLPTKDELLFVHSAGSEFMSALEKAVAKINGCYEALSGATTE 202
DB 1 FMQGEVWEVVDRLPTKDELLFVHSAGSEFMSALEKAVAKINGCYEALSGATTE 60
QY 203 GFEDFTGIAEMYLKRPNNLKKIIOKALQKSLGCSIDTISADSEAVTYQKLVKH 262
DB 61 GFEDFTGIAEMYLKRPNNLKKIIOKALQKSLGCSIDTISADSEAVTYQKLVKH 120
QY 263 AYVTGAEEVSNGLQKLIIRNPMGEVETGRANDNCPSMNTIDPEERERLTPRHEDG 322
DB 121 AYVTGAEEVSNGLQKLIIRNPMGEVETGRANDNCPSMNTIDPEERERLTPRHEDG 180
QY 323 EFWMSFSDLRHSLEICNLTPDLTSDTYKKMKLTAMDGNRRGSTAGCCRNYPNTFW 382
DB 181 EFWMSFSDLRHSLEICNLTPDLTSDTYKKMKLTAMDGNRRGSTAGCCRNYPNTFW 240
QY 383 MNPQYLKLEEDDEEDGSGCTFLVGLIOKRRROKMGEDHTIGFGLIYEPPELISG 442
DB 241 MNPQYLKLEEDDEEDGSGCTFLVGLIOKRRROKMGEDHTIGFGLIYEPPELISG 300
QY 443 QTNHLSKNFLLNTRARSDTFINLREVLNFKLPGEYILVSTFEPNKGDCIFCIRVF 502
DB 301 QTNHLSKNFLLNTRARSDTFINLREVLNFKLPGEYILVSTFEPNKGDCIFCIRVF 360
QY 503 SEKKADYQAVDEIEANLSEEDIDDOVRLLFQVLAGEDAEISAFELQTLIRRLVA 562
DB 361 SEKKADYQAVDEIEANLSEEDIDDOVRLLFQVLAGEDAEISAFELQTLIRRLVA 420
QY 563 K 563
DB 421 K 421

Search completed: July 24, 2003, 12:56:57
Job time : 89.5929 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:28:44 ; Search time 19.3504 Seconds

(without alignments)
1701.191 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713

Sequence: 1 MAGIAAKAKDREAREGLGS.....PENTGTIELDLISWLCPSVL 700

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3704	99.8	700	1	CAN2_HUMAN
2	3522	94.9	700	1	CAN2_MOUSE
3	3518	94.7	700	1	CAN2_RAT
4	3098	83.4	700	1	CAN2_CHICK
5	2579.5	69.5	705	1	CAN2_CHICK
6	2458.5	66.2	714	1	CAN1_HUMAN
7	2444.5	65.8	714	1	CAN1_MOUSE
8	2435	65.6	713	1	CAN1_PIG
9	2427	65.4	713	1	CAN1_RAT
10	2153	58.0	422	1	CAN2_RABIT
11	2014	54.2	702	1	CAN2_HUMAN
12	1897	51.1	810	1	CAN2_CHICK
13	1877.5	50.6	821	1	CAN2_MOUSE
14	1869.5	50.4	821	1	CAN2_RAT
15	1862.5	50.2	821	1	CAN2_HUMAN
16	1619	43.6	324	1	CAN2_PIG
17	1594.5	43.7	828	1	CAN1_DROME
18	1253	33.7	783	1	YKR2_CAEEL
19	1202.5	32.4	758	1	CAN1_SCHNA
20	965.5	25.8	640	1	CAN2_MOUSE
21	957	25.0	207	1	CAN2_BOVIN
22	932.5	25.1	640	1	CAN5_HUMAN
23	931	25.1	302	1	CAN1_HUMAN
24	925.5	24.9	302	1	CAN1_RABIT
25	790	21.3	374	1	CAN3_PIG
26	788	21.2	641	1	CAN6_HUMAN
27	780	21.0	641	1	CAN6_MOUSE
28	772	20.8	641	1	CAN6_RAT
29	634	17.1	666	1	CANA_RAT
30	627.5	16.9	666	1	CANA_MOUSE
31	609	16.4	653	1	CANA_MACFA
32	604	16.3	672	1	CANA_HUMAN
33	458	12.3	263	1	CAN5_BOVIN

34	458	12.3	266	1	CAN5_PIG	P04574 sus scrofa
35	454	12.2	266	1	CAN5_RABIT	P06813 oryctolagus
36	453	12.2	269	1	CAN5_MOUSE	P08456 mus musculus
37	450	12.1	268	1	CAN5_HUMAN	P04632 homo sapien
38	445	12.0	266	1	CAN5_RAT	P04537 ratu
39	426	11.5	1597	1	SOL_DROME	P27388 drosophila
40	240	6.5	198	1	SORC_HUMAN	P30626 homo sapien
41	239	6.4	217	1	GRAN_HUMAN	P28676 homo sapien
42	236	6.4	198	1	SORC_CRITO	P05044 cricetus
43	195	5.3	116	1	CAN3_BOVIN	P51186 bos taurus
44	178	4.8	171	1	SORC_SCHNA	P04743 schistosoma
45	170	4.6	191	1	PCDE_MOUSE	P12815 mus musculus

ALIGNMENTS

RESULT 1

ID	CAN2_HUMAN	STANDARD;	PRT;	700 AA.
AC	P17655; Q16738; Q8WU26; Q9HBB1;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Calpain 2, large [catalytic] subunit precursor (EC 3.4.22.17)			
DE	(Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)			
DE	(Molecular-calpain) (Calpain large polypeptide I2)			
GN	CANP2 OR CANP12			
OS	Homo sapiens (human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8916474; PubMed=2852952;			
RA	Imajo S., Aoki K., Ono S., Emori Y., Kawasaki H., Sugihara H.,			
RA	Suzuki K.;			
RT	"Molecular cloning of the cDNA for the large subunit of the			
RT	high-Ca2+-requiring form of human Ca2+-activated neutral protease.";			
RL	Biochemistry 27:8122-8128(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,			
RA	Datchenko L., Matushita K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Feley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,			
RA	Schuerch A., Schein J.E., Jones S.J.M., Maira M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RL	[4]			

RP SEQUENCE OF 1-79 FROM N.A.
 RC TISSUE=Lymph node;
 RX MEDLINE=69197947; PubMed=2539381;
 RA Hata A., Ohno S., Akita Y., Suzuki K.;
 RT "Randomly reiterate negative enhancer-like elements regulate
 RT transcription of a human gene for the large subunit of calcium-
 RT dependent protease.";
 RL J. Biol. Chem. 264:6404-6411(1989).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX MEDLINE=20105516; PubMed=10639123;
 RA Strobl S., Fernandez-Catalan C., Braun M., Huber R., Masumoto H.,
 RA Nakagawa K., Irie A., Sorimachi H., Bourenkow G., Bartunik H.,
 RA Suzuki K., Bode W.;
 RT "The crystal structure of calcium-free human m-calpain suggests an
 RT electrostatic switch mechanism for activation by calcium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:588-592(2000).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Met-|-Xaa or
 CC Arg-|-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M23254; AAA35645.1; -;
 DR EMBL; AF261089; AAF99682.1; -;
 DR EMBL; BC021303; AAH21303.1; -;
 DR EMBL; J04700; AAA52760.1; -;
 DR PIR; S10590; C1HUH2.
 DR PDB; 1KFU; 07-DEC-01.
 DR PDB; 1KFX; 07-DEC-01.
 DR MEROPS; C02.002; -;
 DR Genew; HGNC:1479; CAPN2.
 DR MIM; 114230; -;
 DR GO; GO:0008234; F:cysteine-type peptidase activity; TAS.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR00169; SHprot_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PRO0704; CALPAIN_C2; 1.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KW Hydrolyase; Thiol protease; Calcium-binding; Repeat; Multigene family;
 KW 3D-structure; Polymorphism.
 FT PROPEP 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
 FT CHAIN 20 700 CALPAIN 2, LARGE (CATALYTIC) SUBUNIT.
 FT DOMAIN 27 355 CALPAIN.
 FT DOMAIN 356 514 CALPAIN.
 FT DOMAIN 515 529 LINKER.
 FT DOMAIN 530 700 LINKER.
 FT CA_BIND 541 552 DOMAIN IV.
 FT EF_HAND 1.

FT CA_BIND 585 596
 FT CA_BIND 615 626
 FT DOMAIN 650 661
 FT DOMAIN 680 691
 FT ACT_SITE 105 105
 FT ACT_SITE 262 262
 FT ACT_SITE 286 286
 FT VARIANT 22 22
 FT VARIANT 568 568
 FT CONFLICT 68 68
 FT CONFLICT 73 74
 FT CONFLICT 256 256
 FT CONFLICT 300 300
 FT CONFLICT 534 534
 FT HELIX 4 16
 FT TURN 17 19
 FT TURN 22 23
 FT STRAND 25 26
 FT HELIX 32 42
 FT TURN 43 43
 FT TURN 55 58
 FT TURN 69 70
 FT STRAND 73 76
 FT HELIX 78 81
 FT STRAND 86 86
 FT STRAND 97 98
 FT TURN 101 102
 FT HELIX 105 113
 FT TURN 114 115
 FT HELIX 118 121
 FT TURN 122 124
 FT TURN 136 137
 FT STRAND 138 144
 FT STRAND 149 155
 FT STRAND 158 161
 FT TURN 162 163
 FT STRAND 164 165
 FT STRAND 169 170
 FT STRAND 175 175
 FT HELIX 177 187
 FT TURN 188 189
 FT HELIX 192 194
 FT TURN 196 197
 FT HELIX 200 206
 FT TURN 207 208
 FT STRAND 211 216
 FT TURN 217 218
 FT TURN 222 223
 FT HELIX 224 232
 FT STRAND 233 235
 FT STRAND 237 241
 FT TURN 253 254
 FT STRAND 264 272
 FT TURN 275 276
 FT STRAND 279 285
 FT TURN 287 288
 FT TURN 303 304
 FT HELIX 310 313
 FT TURN 314 317
 FT STRAND 327 327
 FT HELIX 328 332
 FT STRAND 336 341
 FT HELIX 344 346
 FT TURN 351 352
 FT STRAND 357 357
 FT STRAND 363 365
 FT TURN 367 370
 FT TURN 375 376
 FT TURN 378 380
 FT HELIX 381 383

EF-HAND 2.
 EF-HAND 3.
 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 E -> D (IN OBSNP:25655).
 K -> Q (IN OBSNP:17599).
 /FTID=VAR 014435.
 /FTID=VAR 014436.
 S -> G (IN REF. 4).
 IE -> WR (IN REF. 1).
 O -> K (IN REF. 2).
 N -> S (IN REF. 2).
 V -> F (IN REF. 3).

Query Match 99.8%; Score 3704; DB 1; Length 700;
Best Local Similarity 99.7%; Pred. No. 1,1e-22;
Matches 699; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGIAKLADREAREAGISHERAIXKYNODYALRNECEAGTLPODSEFPALPSALGF 60
DB 1 MAGIAKLADREAREAGISHERAIXKYNODYALRNECEAGTLPODSEFPALPSALGF 60
QY 61 KELGPYSSKTRGRMKRPTICADPOFIIGATRTDICOGALGDCWLLAASLTNBEI 120
DB 61 KELGPYSSKTRGRMKRPTICADPOFIIGATRTDICOGALGDCWLLAASLTNBEI 120
QY 121 LARAVPLNOSFOENYAGIFHFQWQGEWVEVVVDLPTKDGELLFVHSAEGSEFWAL 180
DB 121 LARAVPLNOSFOENYAGIFHFQWQGEWVEVVVDLPTKDGELLFVHSAEGSEFWAL 180
QY 181 LERAVKINCCYALSGGATTEGPEDEPTGIAEMWELEKRPPLFKIIOKALQKSLGC 240
DB 181 LERAVKINCCYALSGGATTEGPEDEPTGIAEMWELEKRPPLFKIIOKALQKSLGC 240
QY 241 SIDITSAADSEAITFOKLVGHAYSVTGAEEVNSGSLQKILRPNMGVEWTRMND 300
DB 241 SIDITSAADSEAITFOKLVGHAYSVTGAEEVNSGSLQKILRPNMGVEWTRMND 300
QY 301 CPSEWNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPTDITSDIYKKWGLTX 360
DB 301 CPSEWNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPTDITSDIYKKWGLTX 360
QY 361 MOGNMRGSGTAGCCRYNPTFMNPPQTLLEEDDEDEEGEGCTVLVGLIOKRROR 420
DB 361 MOGNMRGSGTAGCCRYNPTFMNPPQTLLEEDDEDEEGEGCTVLVGLIOKRROR 420
QY 421 KMGEDMHTTIFGTYEVEBELSGQTNHLSKNPFLTPRARSRPTFNLREVLNRFKLPQ 480
DB 421 KMGEDMHTTIFGTYEVEBELSGQTNHLSKNPFLTPRARSRPTFNLREVLNRFKLPQ 480
QY 481 EYLLVSTEEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDISDDIDGVRRLFAQ 540
DB 481 EYLLVSTEEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDISDDIDGVRRLFAQ 540
QY 541 LAEDEDEISAFELQTLIRRLARODIKSGFSIFETKINWMLDSGSGKLKERYIL 600
DB 541 LAEDEDEISAFELQTLIRRLARODIKSGFSIFETKINWMLDSGSGKLKERYIL 600
QY 601 WTKIOYKOKYREIDVDRSGTMSYEMRKALBEGAFMPCQLHOVIVARPADQLIIDPD 660
DB 601 WTKIOYKOKYREIDVDRSGTMSYEMRKALBEGAFMPCQLHOVIVARPADQLIIDPD 660
QY 661 NFVRCIVRLLETLFKIFKQIDPENTGIIELDLISWLCPSVL 700
DB 661 NFVRCIVRLLETLFKIFKQIDPENTGIIELDLISWLCPSVL 700

RA Dear T.N., Matena K., Vingron M., Boehm T.;
RT "A new subfamily of vertebrate calpains lacking a calmodulin-like
RL domain: implications for calpain regulation and evolution.",
RN Genomics 45:175-184 (1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Ozaki Y.;
RN Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA TISSUE=CNS;
RA Glaes J.D., Nash N.R., Dry I., Culver D., Wesselingh S.;
RT "Cloning of m-calpain from mouse nervous system";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyzes limited proteolysis of substrates involved in
CC cytoskeletal remodeling and signal transduction (By similarity).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
CC Arg-Xaa with Leu or Val as the p2 residue.
CC -1- COFACTOR: Binds 3 calcium ions.
CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
CC of calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding.
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).

CC EMBL; Y10139; CA471227.1; -;
CC EMBL; D38117; BAA32964.1; -;
CC EMBL; AF015038; AAB94029.1; -;
CC HSSP; Q07009; IDFO.
CC MGD; MG1:88264; Capn2.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease_C2.
DR InterPro; IPR000169; SHPoc_acstite.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00036; efhand; 2.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyfepc; 1.
DR SMART; SM00054; Efh; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR PROSITE; PS00640; THIOL_PROTEASE ASN; FALSE NEG.
KW Hydrolyase; Thiol protease; Calcium-binding; Repeat; Multisene family.
FT CHAIN 1 19
FT DOMAIN 20 700
FT DOMAIN 20 355
FT DOMAIN 356 514
FT DOMAIN 515 528
FT DOMAIN 530 700
FT CA_BIND 541 552
FT CA_BIND 585 596
FT CA_BIND 615 626
FT DOMAIN 650 661
FT DOMAIN 680 691
FT ACT_SITE 105 109
FT ACT_SITE 262 262
FT ACT_SITE 286 286
BY SIMILARITY.

FT CONFLICT 194 194 A -> T (IN REF. 1).
 FT CONFLICT 212 212 A -> G (IN REF. 2).
 FT CONFLICT 402 402 E -> G (IN REF. 1).
 SQ SEQUENCE 700 AA; 79871 MW; 682146B20968316 CRC64;

Query Match 94.9%; Score 3522; DB 1; Length 700;
 Best Local Similarity 93.6%; Pred. No. 7,1e-221;
 Matches 655; Conservative 24; Mismatches 21; Indels 0; Gaps 0;

QY 1 MAGAATLADKREABEGSHERRAKYLNQDYALRHCACGLTLPDPSFPAISLGF 60
 DB 1 MAGAATLADKREABEGSHERRAKYLNQDYALRHCACGLTLPDPSFPAISLGF 60
 QY 61 KELGYSKTRGIMKRPTEICADPOFIIIGATRTDICOGLADGCMWLAALASITLNEBI 120
 DB 61 KELGYSKTRGIMKRPTEICADPOFIIIGATRTDICOGLADGCMWLAALASITLNEBI 120
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGBWVVDRLPYOGELLFVHSAEGSEFWSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGBWVVDRLPYOGELLFVHSAEGSEFWSAL 180
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGBWVVDRLPYOGELLFVHSAEGSEFWSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGBWVVDRLPYOGELLFVHSAEGSEFWSAL 180
 QY 181 LEKAVAKINGCYALSGATTEGFEPTGIAEWEYLKPPPNLFKIIQKALOKSLIGC 240
 DB 181 LEKAVAKINGCYALSGATTEGFEPTGIAEWEYLKPPPNLFKIIQKALOKSLIGC 240
 QY 241 SIDITSADSEAITFOKLVKHAASVYGAEEVNSGLQKLIIRINPGEVEMTGRMNDN 300
 DB 241 SIDITSADSEAITFOKLVKHAASVYGAEEVNSGLQKLIIRINPGEVEMTGRMNDN 300
 QY 301 CPSWNTIDPEERERLTRHEDDEFWMSFDFLRHYSRLICNTLPDITLSTPYKKMLTK 360
 DB 301 CPSWNTIDPEERERLTRHEDDEFWMSFDFLRHYSRLICNTLPDITLSTPYKKMLTK 360
 QY 301 CPSWNTIDPEERERLTRHEDDEFWMSFDFLRHYSRLICNTLPDITLSTPYKKMLTK 360
 DB 301 CPSWNTIDPEERERLTRHEDDEFWMSFDFLRHYSRLICNTLPDITLSTPYKKMLTK 360
 QY 361 MDGWRGSGTAGGCRNYPNTFWMNPQYILKLEEDDEDEDEDECTFVGLIQHRROR 420
 DB 361 MDGWRGSGTAGGCRNYPNTFWMNPQYILKLEEDDEDEDEDECTFVGLIQHRROR 420
 QY 421 KMGEDMTITGFIYEVEBELSGQTNHLSKNFFLTNRARESDPFINLREVLNFKLP 480
 DB 421 KMGEDMTITGFIYEVEBELSGQTNHLSKNFFLTNRARESDPFINLREVLNFKLP 480
 QY 481 EYIVPSTFEPNKGDFCIRVSEKADYQAVDEIEANIEEDIDGGRRLFVQ 540
 DB 481 EYIVPSTFEPNKGDFCIRVSEKADYQAVDEIEANIEEDIDGGRRLFVQ 540
 QY 541 LAGSDAISIAPFELTILRLVAKRODIKSDGSIETCKIMWMDLDEDSGLGKEFYIL 600
 DB 541 LAGSDAISIAPFELTILRLVAKRODIKSDGSIETCKIMWMDLDEDSGLGKEFYIL 600
 QY 601 WTIKIOYKTYREIDVDSGTMSYEMKALEAGFKMPCQHQVYVARFADDLIIDFD 660
 DB 601 WTIKIOYKTYREIDVDSGTMSYEMKALEAGFKMPCQHQVYVARFADDLIIDFD 660
 QY 661 NFVRCIVRLFTLFKIFQOLDPENTGTELLIISWLCPSVL 700
 DB 661 NFVRCIVRLFTLFKIFQOLDPENTGTELLIISWLCPSVL 700

RESULT 3
 CAN2 RAT STANDARD; PRG: 700 AA.
 AC 007009;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 2, large [catalytic] subunit precursor (EC 3.4.22.17)
 DE [calcium-activated neutral proteinase] (CAMP) (M-type) (M-calpain)
 DE [Millemolier-calpain].
 GN CAPN2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94032492; PubMed=8218419;
 RA Deluca C.I., Davies P.L., Samis J.A., Elce J.S.;
 RT "Molecular cloning and bacterial expression of cDNA for rat calpain
 RT II 80 kDa subunit.";
 RL Biochim. Biophys. Acta 1216:81-93(1993).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=21240297; PubMed=11342050;
 RA Moldoveanu T., Hosfield C.M., Jia Z., Elce J.S., Davies P.L.;
 RT "Ca(2+)-induced structural changes in rat m-calpain revealed by
 RT partial proteolysis.";
 RL Biochim. Biophys. Acta 1545:245-254(2001).
 RN [3]
 RP MEDLINE=21269273; PubMed=11102442;
 RX Hosfield C.M., Moldoveanu T., Davies P.L., Elce J.S., Jia Z.;
 RT "Calpain mutants with increased Ca2+ sensitivity and implications for
 RT the role of the C(2)-like domain.";
 RL J. Biol. Chem. 276:7404-7407(2001).
 RN [4]
 RP MUTAGENESIS OF CYS-105, HIS-262, ASN-286 AND TRP-288.
 RX MEDLINE=95361909; PubMed=7635186;
 RA Arthur J.S., Gauthier S., Elce J.S.;
 RT "Active site residues in m-calpain: identification by site-directed
 RT mutagenesis.";
 RL FEBS Lett. 368:397-400(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=2069318; PubMed=10601010;
 RA Hosfield C.M., Elce J.S., Davies P.L., Jia Z.;
 RT "Crystal structure of calpain reveals the structural basis for
 RT Ca(2+)-dependent protease activity and a novel mode of enzyme
 RT activation.";
 RL EMBO J. 18:6880-6889(1999).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or
 CC Arg|-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L09120; AAA16327.1; -.
 DR PIR; S38361; S38361.
 DR PDB; 1DFO; 26-NOV-01.
 DR MEROPS; C02.002; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHPTC_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR SMART; SM00054; Efn; 2.

DR PROSITE; PS00018; EF_HAND. 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS. 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
 KM Hydroxylase; Thiol protease; Calcium-binding; Repeat; Multigene family;
 3D-structure.
 FT PROPER 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
 FT CHAIN 20 700 CALPAIN 2, LARGE [CATALYTIC] SUBUNIT.
 FT DOMAIN 20 355 CALPAIN.
 FT DOMAIN 356 514 DOMAIN III.
 FT DOMAIN 515 529 LINKER.
 FT DOMAIN 530 700 DOMAIN IV.
 FT CA_BIND 541 552 EF-HAND 1.
 FT CA_BIND 585 596 EF-HAND 2.
 FT CA_BIND 615 626 EF-HAND 3.
 FT DOMAIN 650 661 ANCESTRAL CALCIUM SITE 4.
 FT DOMAIN 680 691 ANCESTRAL CALCIUM SITE 5.
 FT ACT_SITE 105 105
 FT ACT_SITE 262 262
 FT ACT_SITE 286 286
 FT MUTAGEN 105 105 C-S: LOSS OF ACTIVITY.
 FT MUTAGEN 226 226 K-S: DECREASE OF 12% OF THE ACTIVITY.
 FT MUTAGEN 230 230 K-S: NO EFFECT.
 FT MUTAGEN 230 230 K-S: DECREASE OF 84% OF THE ACTIVITY.
 FT MUTAGEN 234 234 K-S: DECREASE OF 85% OF THE ACTIVITY.
 FT MUTAGEN 234 234 K-S: DECREASE OF 20% OF THE ACTIVITY.
 FT MUTAGEN 262 262 H-S: LOSS OF ACTIVITY.
 FT MUTAGEN 286 286 N-S: LOSS OF ACTIVITY.
 FT MUTAGEN 288 288 W-S: DECREASE OF 95% OF THE ACTIVITY.
 FT MUTAGEN 504 504 E-S: DECREASE OF 10% OF THE ACTIVITY.
 FT MUTAGEN 504 504
 SQ SEQUENCE 700 AA; 79919 MW; 29680DC3BEE5B90 CRC64;
 Query Match 94.7%; Score 3518; DB 1; Length 700;
 Best Local Similarity 93.4%; Pred. No. 1,3e-220;
 Matches 654; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

QY 1 MAGIAKLADRAAEGLSGHERAIKYNODYALNECTIEAGTLFQDPSPFPAISALGF 60
 DB 1 MAGIAKLADRAAEGLSGHERAIKYNODYALNECTIEAGTLFQDPSPFPAISALGF 60
 QY 61 KEIGVSSKTRGMRWRPTICADPORIIGATRTDICQALDPCWLLAASLTNEET 120
 DB 61 KEIGVSSKTRGMRWRPTICADPORIIGATRTDICQALDPCWLLAASLTNEET 120
 QY 121 LAIVPLNDSGFQENYAGIFEFQWQGEWYEVVDRLPTKDGELFVHSAEGSEFWSAL 180
 DB 121 LAIVPLNDSGFQENYAGIFEFQWQGEWYEVVDRLPTKDGELFVHSAEGSEFWSAL 180
 QY 121 LAIVPLNDSGFQENYAGIFEFQWQGEWYEVVDRLPTKDGELFVHSAEGSEFWSAL 180
 DB 121 LAIVPLNDSGFQENYAGIFEFQWQGEWYEVVDRLPTKDGELFVHSAEGSEFWSAL 180
 QY 181 LEKAVAKINGCYALSGATTEGPEDEFTGIAEWELKRPPLFKIIQKALQKSLIGC 240
 DB 181 LEKAVAKINGCYALSGATTEGPEDEFTGIAEWELKRPPLFKIIQKALQKSLIGC 240
 QY 181 LEKAVAKINGCYALSGATTEGPEDEFTGIAEWELKRPPLFKIIQKALQKSLIGC 240
 DB 181 LEKAVAKINGCYALSGATTEGPEDEFTGIAEWELKRPPLFKIIQKALQKSLIGC 240
 QY 241 SIDITSAADSEALTTPQKLVGHAYSVTGAEEVNSGLQKLIIRNPMGEVETGRMND 300
 DB 241 SIDITSAADSEALTTPQKLVGHAYSVTGAEEVNSGLQKLIIRNPMGEVETGRMND 300
 QY 241 SIDITSAADSEALTTPQKLVGHAYSVTGAEEVNSGLQKLIIRNPMGEVETGRMND 300
 DB 241 SIDITSAADSEALTTPQKLVGHAYSVTGAEEVNSGLQKLIIRNPMGEVETGRMND 300
 QY 301 CPSEMTIDPEERERLFRHEDGEFMMSPDFLRHYSLEICNLTPTDITSDYKKMKLTK 360
 DB 301 CPSEMTIDPEERERLFRHEDGEFMMSPDFLRHYSLEICNLTPTDITSDYKKMKLTK 360
 QY 301 CPSEMTIDPEERERLFRHEDGEFMMSPDFLRHYSLEICNLTPTDITSDYKKMKLTK 360
 DB 301 CPSEMTIDPEERERLFRHEDGEFMMSPDFLRHYSLEICNLTPTDITSDYKKMKLTK 360
 QY 361 MDGNMGRSTAGSCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
 DB 361 MDGNMGRSTAGSCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
 QY 361 MDGNMGRSTAGSCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
 DB 361 MDGNMGRSTAGSCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
 QY 421 KMGEDEHTTGFGLYVEPEELSGQTNHLSGNFPLTNARARSSTPFLNREVLNRFKLPFG 480
 DB 421 KMGEDEHTTGFGLYVEPEELSGQTNHLSGNFPLTNARARSSTPFLNREVLNRFKLPFG 480
 QY 481 EYLVSTPEPNKDGPCIRVFSEKADYQAVDDEIENAEEDPDDIDDPVRRFAQ 540
 DB 481 EYLVSTPEPNKDGPCIRVFSEKADYQAVDDEIENAEEDPDDIDDPVRRFAQ 540
 QY 481 EYLVSTPEPNKDGPCIRVFSEKADYQAVDDEIENAEEDPDDIDDPVRRFAQ 540
 DB 481 EYLVSTPEPNKDGPCIRVFSEKADYQAVDDEIENAEEDPDDIDDPVRRFAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLARODIKSDGFSIETCKIWMIDSDSGKLGKKEFYIL 600
 DB 541 LAGEDAEISAFELQTLIRVLARODIKSDGFSIETCKIWMIDSDSGKLGKKEFYIL 600

DB 541 LAGEDAEISAFELQTLIRVLARODIKSDGFSIETCKIWMIDSDSGKLGKKEFYIL 600
 QY 601 WTIOYOKYRRIIDVDSGTNYSYMRKALEAGFMPCQHQVVARADQILIDPD 660
 DB 601 WTIOYOKYRRIIDVDSGTNYSYMRKALEAGFMPCQHQVVARADQILIDPD 660
 QY 661 NFVRCVRLLETLFKIFKQDPDENTGTIELDLSWLCFSYL 700
 DB 661 NFVRCVRLLETLFKIFKQDPDENTGTIELDLSWLCFSYL 700
 DB 661 NFVRCVRLLETLFKIFKQDPDENTGTIELDLSWLCFSYL 700
 DB 661 NFVRCVRLLETLFKIFKQDPDENTGTIELDLSWLCFSYL 700

RESULT 4
 ID CAN2 CHICK STANDARD; PRT; 700 AA.
 AC 092178;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 2, large [catalytic] subunit precursor (EC 3.4.22.17)
 DE (Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)
 DE (Molluscal-calpain).
 GN CANP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=95260862; PubMed=7742367;
 RA Sorinachi H., Teukahara T., Okada-Ban M., Sugita H., Ishiura S.,
 RA Suzuki K.,
 RT "Identification of a third ubiquitous calpain species -- chicken
 muscle expresses four distinct calpains".
 RL Biochim. Biophys. Acta 1261:381-393(1995).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction (By similarity).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or
 CC Arg|-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).

CC EMBL; D38026; BAA07228.1; -
 CC PIR; S57194; S57194.
 CC HSP; Q07009; 1DFO.
 CC MEROPS; C02.002; -
 CC InterPro; IPR002048; EF-hand.
 CC InterPro; IPR001300; Protease_C2.
 CC InterPro; IPR000169; SHPOT_acetate.
 CC Pfam; PF01067; Calpain III; 1.
 CC Pfam; PF000036; ehand_3.
 CC Pfam; PF00648; Peptidase_C2; 1.
 CC PRINTS; PR00704; CALPAIN.
 CC SMART; SM00720; calpain_III; 1.
 CC SMART; SM00230; Cyspc; 1.

DR SMART; SMO0054; Efn, 3.
 DR PROSITE; PS00018; EF HAND; 2.
 DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOI_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOI_PROTEASE_ASN; FALSE NEG.
 KW Hydroxylase; thiol protease; Calcium-binding; Multigene family.
 FT PROPEP 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
 FT CHAIN 20 700 CALPAIN 2, LARGE (CATALYTIC) SUBUNIT.
 FT DOMAIN 20 355 CALPAIN.
 FT DOMAIN 356 514 CALPAIN III.
 FT DOMAIN 515 529 LINKER.
 FT DOMAIN 530 700 DOMAIN IV.
 FT CA_BIND 541 552 EF-HAND 1.
 FT CA_BIND 585 596 EF-HAND 2.
 FT CA_BIND 615 626 EF-HAND 3.
 FT DOMAIN 650 661 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT ACT_SITE 680 691 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT_SITE 105 105 BY SIMILARITY.
 FT ACT_SITE 262 262 BY SIMILARITY.
 FT ACT_SITE 286 286 BY SIMILARITY.
 SQ SEQUENCE 700 AA; 79228 MW; C3AEDB39CCB56D3B CRC64;

Query Match 83.4%; Score 3098; DB 1; Length 700;
 Best Local Similarity 81.1%; Pred. No. 2.2e-193;
 Matches 568; Conservative 67; Mismatches 65; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKRDEAEGLSHERAIKYLNODYEALRNCELAGTLPDPSPFALPSALGF 60
 1 MAGMAALAKRRAAAGAGRGQAVPYIGDFGALRRCLOGSLFDFSPSPAGPALGY 60
 DB 61 RELDPNSYKTKGVWVCRTPELSCPRFIAGATRTDIOGALGCMWLAALASLTNEEI 120
 QY 61 KELGPYSSKTRGMKRPTEICADPOFIIGATRTDIOGALGCMWLAALASLTNEEI 120
 DB 61 RELDPNSYKTKGVWVCRTPELSCPRFIAGATRTDIOGALGCMWLAALASLTNEEI 120
 QY 121 LARVPLNQSQENYAGIFHFQFQYQGWVWVDDRLPTDGGELPHFASGESEPSAL 180
 DB 121 LARVPLNQSQENYAGIFHFQFQYQGWVWVDDRLPTDGGELPHFASGESEPSAL 180
 QY 121 LARVPLNQSQENYAGIFHFQFQYQGWVWVDDRLPTDGGELPHFASGESEPSAL 180
 DB 121 LARVPLNQSQENYAGIFHFQFQYQGWVWVDDRLPTDGGELPHFASGESEPSAL 180
 QY 181 LEKAVAKINGCYEALSGATGEGFDFGIAEWYELKKPPNFKIIOKALOGSLLGC 240
 DB 181 LEKAVAKINGCYEALSGATGEGFDFGIAEWYELKKPPNFKIIOKALOGSLLGC 240
 QY 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRNND 300
 DB 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRNND 300
 QY 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRNND 300
 DB 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRNND 300
 QY 301 CPNMTWIDPDERERKLRHEDGEMFMSDYLKHSLEICNLTPDTITSTYTKKMLTK 360
 DB 301 CPNMTWIDPDERERKLRHEDGEMFMSDYLKHSLEICNLTPDTITSTYTKKMLTK 360
 QY 301 CPNMTWIDPDERERKLRHEDGEMFMSDYLKHSLEICNLTPDTITSTYTKKMLTK 360
 DB 301 CPNMTWIDPDERERKLRHEDGEMFMSDYLKHSLEICNLTPDTITSTYTKKMLTK 360
 QY 361 MDGMMRGSTAGGCRNYPNTFPMNPQVYIKLEEDDEDEDESGGTFVGLIOKRRROR 420
 DB 361 MDGMMRGSTAGGCRNYPNTFPMNPQVYIKLEEDDEDEDESGGTFVGLIOKRRROR 420
 QY 361 MDGMMRGSTAGGCRNYPNTFPMNPQVYIKLEEDDEDEDESGGTFVGLIOKRRROR 420
 DB 361 MDGMMRGSTAGGCRNYPNTFPMNPQVYIKLEEDDEDEDESGGTFVGLIOKRRROR 420
 QY 421 KMGDMHTIGFIVPEBELSGQTNILSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 DB 421 KMGDMHTIGFIVPEBELSGQTNILSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 QY 421 KMGDMHTIGFIVPEBELSGQTNILSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 DB 421 KMGDMHTIGFIVPEBELSGQTNILSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 QY 481 EYILVPSFEENKGGPCIRVSEKKADYQAVDEIEANLEPFISEDDIDDGVARLPAQ 540
 DB 481 EYILVPSFEENKGGPCIRVSEKKADYQAVDEIEANLEPFISEDDIDDGVARLPAQ 540
 QY 481 EYILVPSFEENKGGPCIRVSEKKADYQAVDEIEANLEPFISEDDIDDGVARLPAQ 540
 DB 481 EYILVPSFEENKGGPCIRVSEKKADYQAVDEIEANLEPFISEDDIDDGVARLPAQ 540
 QY 541 LAGEDAETSAPELOTLIRVLAKRODISDGSFIETCIWMDLSDSGKLGKEFYTL 600
 DB 541 LAGEDAETSAPELOTLIRVLAKRODISDGSFIETCIWMDLSDSGKLGKEFYTL 600
 QY 541 LAGEDAETSAPELOTLIRVLAKRODISDGSFIETCIWMDLSDSGKLGKEFYTL 600
 DB 541 LAGEDAETSAPELOTLIRVLAKRODISDGSFIETCIWMDLSDSGKLGKEFYTL 600
 QY 601 WTAKIOKQIYREIDVDSGTNNSYEMKALEAGFKMPCOLHQYVAFADQLIIDFD 660
 DB 601 WTAKIOKQIYREIDVDSGTNNSYEMKALEAGFKMPCOLHQYVAFADQLIIDFD 660
 QY 601 WTAKIOKQIYREIDVDSGTNNSYEMKALEAGFKMPCOLHQYVAFADQLIIDFD 660
 DB 601 WTAKIOKQIYREIDVDSGTNNSYEMKALEAGFKMPCOLHQYVAFADQLIIDFD 660
 QY 661 NPYRCLVLETLFKFKOLDPENGTITLDLISWICSVL 700
 DB 661 NPYRCLVLETLFKFKOLDPENGTITLDLISWICSVL 700
 QY 661 NPYRCLVLETLFKFKOLDPENGTITLDLISWICSVL 700
 DB 661 NPYRCLVLETLFKFKOLDPENGTITLDLISWICSVL 700

RESULT 5
 ID CANX_CHICK STANDARD; PRT; 705 AA.
 AC P00789;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain, large (catalytic) subunit (EC 3.4.22.17) (calcium-
 DE activated neutral proteinase) (CAMP) (Mu/M-type).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 CX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85061606; PubMed=6095110;
 RA Ohno S., Emori Y., Imaizumi S., Kawasaki H., Kishimoto A., Suzuki K.;
 RT "Evolutionary origin of a calcium-dependent protease by fusion of
 RT genes for a thiol protease and a calcium-binding protein";
 RL Nature 312:566-570(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86082358; PubMed=3000828;
 RA Emori Y., Ohno S., Tobita M., Suzuki K.;
 RT "Gene structure of calcium-dependent protease retains the ancestral
 RT organization of the calcium-binding protein gene";
 RL FEBS Lett. 194:249-252(1986).
 RN [3]
 RP CALCIUM-BINDING DATA.
 RX MEDLINE=87279982; PubMed=3038855;
 RA Minami Y., Emori Y., Kawasaki H., Suzuki K.;
 RT "E-F hand structure-domain of calcium-activated neutral protease
 RT (CAMP) can bind Ca²⁺ ions";
 RL J. Biochem. 101:889-895(1987).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=95260862; PubMed=7742367;
 RA Szymanski H., Tsukahara T., Okada-Ban M., Sugita H., Ishiura S.,
 RA Suzuki K.;
 RT "Identification of a third ubiquitous calpain species -- chicken
 RT muscle expresses four distinct calpains";
 RL Biochem. Biophys. Acta 1261:381-393(1995).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of large (catalytic) and a small (regulatory)
 CC subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca²⁺ binding (By similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -1- CAUTION: THIS PROTEIN WAS PREVIOUSLY THOUGHT TO BE M-CALPAIN BUT
 CC HAS SINCE BEEN FOUND TO BE AN INTERMEDIATE FORM BETWEEN THE M AND
 CC MU TYPES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL, X01415, CAA25658.1; -
 DR F1R; A00979; CICH.
 DR HSP; Q07009; IDFO.
 DR MEROPS; C02.003; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR00169; SH3prot_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR SMART; SM00054; Efn; 2.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
 DR Hydrolase; Thiol protease; Calcium-binding; Multigene family.
 KW DOMAIN; 23 358 CALPAIN.
 FT DOMAIN; 359 517 DOMAIN III.
 FT DOMAIN; 518 533 LINKER.
 FT DOMAIN; 534 704 DOMAIN IV.
 FT CA_BIND; 545 556 EF-HAND 1.
 FT CA_BIND; 589 600 EF-HAND 2.
 FT CA_BIND; 619 630 EF-HAND 3.
 FT DOMAIN; 654 665 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT DOMAIN; 684 695 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT_SITE; 108 108 BY SIMILARITY.
 FT ACT_SITE; 265 265 BY SIMILARITY.
 FT ACT_SITE; 289 289 BY SIMILARITY.
 SQ SEQUENCE 705 AA; 80351 MW; ABCDC65629B48AA CRC64;
 Query Match 69.5%; Score 2579.5; DB 1; Length 705;
 Best Local Similarity 66.6%; Pred. No. 8.9e-160;
 Matches 465; Conservative 111; Mismatches 121; Indels 1; Gaps 1;

QY 542 AGEDAEISAELEOTIIRRVIAKRODIKSDGFSIETCKINWMDLSDGSGKLGIKEFYIIM 601
 DB 546 AGEDMEISVLEKTIINRVIAKRDLDKDFSDSCRNWNLMDKXGSAELGLVEFQILM 605
 QY 602 TKIQXQKYRELDVDRSGMNSYEMRKALEEGFPCQGHQVIYARFADDDLIIDPDN 661
 DB 606 NKIRSWLITFRQYDLKSGTMSSEYEMRMALIESGFLNKLKHQVAVRYADAEITGVDFDN 665
 QY 662 FVACLVLETLTFKIFKQDPENTGTIELDISMLCFSV 699
 DB 666 FVCCLVKLETFMFRFFHSMDRDGIGTAVMLAEMLLITM 703

RESULT 6

CANI_HUMAN STANDARD; PRT; 714 AA.
 ID CANI_HUMAN
 AC P07364,
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated
 DE neutral proteinase) (CANP) (Mu-type) (mucanp) (micromolar-calpain).
 GN CANP1 OR CANP1L.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6301172; PubMed=3017764;
 RA Aoki K., Imaishi S., Ohno S., Emori Y., Koike M., Koseki G., Suzuki K.;
 RT "Complete amino acid sequence of the large subunit of the low-Ca2+-
 RT requiring form of human Ca2+-activated neutral protease (mucanp)
 RT deduced from its cDNA sequence.";
 RL ERBS Lett. 205:313-317(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90380278; PubMed=2400579;
 RA Sorimachi H., Ohmi S., Emori Y., Kawasaki H., Saito T.C., Ohno S.,
 RA Minami Y., Suzuki K.;
 RT "A novel member of the calcium-dependent cysteine protease family.";
 RL Biol. Chem. Hoppe-Seyler 371:171-176(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
 RA Datchenko L., Marnusha K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butcherfield Y.S.N., Krzywiński M.I., Skalska U., Smalls D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Merra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of

CC calcium and inhibited by calpastatin.
 CC -1 SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding (By similarity).
 CC -1 SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04366; CAA27881.1; -
 DR EMBL; BC008751; AAH08751.1; -
 DR PIR; A26213; C1H0H.
 DR HSSP; Q07009; IDFO.
 DR MEROPS; C02.001; -
 DR Genew: HGNC:1476; CAPNL.
 DR MIM; 114220; -
 DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHprot_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3_C2; 1.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR SMART; SM00054; Efb; 3.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KM Hydroxylase; Thiol protease; Calcium-binding; Multigene family.
 FT DOMAIN 30 365
 FT DOMAIN 365 526
 FT DOMAIN 527 542
 FT DOMAIN 543 554
 FT CA_BIND 554 565
 FT CA_BIND 598 609
 FT CA_BIND 628 639
 FT DOMAIN 663 674
 FT DOMAIN 693 704
 FT ACT_SITE 115 115
 FT ACT_SITE 272 272
 FT ACT_SITE 296 296
 FT CONFLICT 548 548 K -> N (IN REF. 3).
 SQ SEQUENCE 714 AA; 81889 MW; 1CB6D7C56D063498 CRC64;

Query Match 66.2%; Score 2458.5; DB 1; Length 714;
 Best Local Similarity 63.1%; Pred. No. 6.3e-152;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

QY 3 GIAKLAKDRRAAGLSGSHERAIKYINDYALNRECEAGTLQODSEFPALPALKGEK 62
 DB 13 GVSAAVOQQRARELGLGHENAIKYLQGYDQLRVCQSGLTFRDEAFPPVPSOLGKD 72
 QY 63 LGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDDCMWLAIAISLITNEELIA 122
 DB 73 LGRNPSKTYGKMKRPTEILSNPOFIYDAGARTDICGALDDCMWLAIAISLITNDILH 132
 QY 123 RVEVLNOSPENYAGIFHFQFQWQYGEWVVDRLPTKDELLEFVSHAGSESEFMSALE 182
 DB 133 RVVPHGSGFQNGYAGIFHFQFQWQYGEWVVDRLPTKDELLEFVSHAGSESEFMSALE 192
 QY 183 KAYAKINGCYALSGGATTESEDFDTGIAWYELKPPPLVFKTIQALOKGSLGCSI 242
 DB 193 KAYAKVNGSEYALSGGATTESEDFDTGIAWYELKPPPLVFKTIQALOKGSLGCSI 252

QY 243 DITSADSEAITFOKLVKGAVSVTGAEVEYNSGLQKILRIRNPNGEWMTGRNDNCP 302
 DB 253 DISSVLMEMAITFKLVKGAVSVTGAKQVNYGQVSLIRNPNNGEWEWMTGAMDSSS 312
 QY 303 SMTIDPEERERLRRHDEGEFMSFSDFIRHYSRLICNLIPDITLSDTYKKWKLTKMD 362
 DB 313 EMMNVDPYERDQLRVKMEDEEFMMSFPDFKREFTRELEICNLTPDALKSRTIRKMTLLYE 372
 QY 363 GMMRSGTAGCCNRYPTFMNPOYLKLEEDNEDE--DSESGTEFLVLQKHRROR 420
 DB 373 GYMRSGTAGCCNRYPTFMNPOYLKLEEDNEDE--DSESGTEFLVLQKHRROR 432
 QY 421 KMGEDMHTTGFYVEPEELSGQNTIHLNKFELTNRRARRSDPTINLRVNLRFKLPG 480
 DB 433 RFGDMHTTGFYVEPEELSGQNTIHLNKFELTNRRARRSDPTINLRVNLRFKLPG 492
 QY 481 EYLVDPSTPEPNKDGFCIRVSEKADYQVADDEIBANL--EFPDISEDIDGVRRLPA 539
 DB 493 EYLVDPSTPEPNKDGFCIRVSEKADYQVADDEIBANL--EFPDISEDIDGVRRLPA 552
 QY 540 QLAGEDEIAFELQTLIRVLAKRQDICKSGFSIETCKIMVMDLSDSGKGLKEFYI 599
 DB 553 QLAGEDEIAFELQTLIRVLAKRQDICKSGFSIETCKIMVMDLSDSGKGLKEFYI 612
 QY 600 LMTKIQKQKTYREIDVDRSGTNNSYEMRKALEEAGFKMPCQHQYIVARFADQILIF 659
 DB 613 LMTKIQKQKTYREIDVDRSGTNNSYEMRKALEEAGFKMPCQHQYIVARFADQILIF 672
 QY 660 DNFVRCIVLETLFKIPKQIDPENTGTIELDISWL 695
 DB 673 DNFVRCIVLETLFKIPKQIDPENTGTIELDISWL 708

RESULT 7
 ID CAN1_PIG STANDARD; PRT; 714 AA.
 AC P35750; Q29600; Q9N0M6;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 1, large (catalytic) subunit (EC 3.4.22.17) (Calcium-activated
 DE neutral proteinase) (CANP) (Mu-type) (mucanp) (micromolar-calpain).
 GN CAPNL.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP Smith T.P.L., Simmen F.A., Vallet J.A.;
 RT "Rapid Communication: nucleotide sequences of two isoforms of porcine
 RT micromolar calcium-activated neutral protease 1 (mucalpain) cDNA.";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE OF 326-415 FROM N.A.
 RP TISSUE=Small intestine;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 528-623 FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=94146155; Pubmed=8312396;
 RA Sun W., Ji S.Q., Ebert P.J., Bitwell C.A., Hancock D.L.;
 RT "Cloning the partial cDNAs of mu-calpain and m-calpain from porcine
 RT skeletal muscle.";
 RL Biochimie 75:931-936(1993).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodelling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the p2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions (By similarity).


```

CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding (by similarity).
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC DR EMBL; AF63610; AAF73444.1; -.
CC DR EMBL; F14611; CAA23154.1; -.
CC DR EMBL; U01180; AAA65125.1; -.
CC DR HSSP; Q07009; IDFO.
CC DR MEROPS; C02.001; -.
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR001300; Protease_C2.
CC DR InterPro; IPR00169; SHPTC_acsite.
CC DR Pfam; PF01067; Calpain_III; 1.
CC DR Pfam; PF00648; Peptidase_C2; 1.
CC DR PRINTS; PR00704; CALPAIN.
CC DR SMART; SM00720; calpain_III; 1.
CC DR SMART; SM00230; Cyspc; 1.
CC DR SMART; SM00054; Efn; 3.
CC DR PROSITE; PS00018; EF_HAND; 2.
CC DR PROSITE; PS00139; THIOL_PROTEASE_HIS; FALSE NEG.
CC DR PROSITE; PS00639; THIOL_PROTEASE_ASY; FALSE NEG.
CC DR PROSITE; PS00640; THIOL_PROTEASE_ASY; FALSE NEG.
CC DR HydroLase; Thiol; protease; Calcium-binding; Multigene family.
CC FT DOMAIN 30 365
CC FT DOMAIN 366 526
CC FT DOMAIN 527 542
CC FT DOMAIN 543 713
CC FT CA_BIND 554 565
CC FT CA_BIND 598 609
CC FT CA_BIND 628 639
CC FT CA_BIND 693 704
CC FT DOMAIN 693 704
CC FT ACT_SITE 115 115
CC FT ACT_SITE 272 272
CC FT ACT_SITE 296 296
CC FT ACT_SITE 528 528
CC FT CONFLICT 531 531
CC FT CONFLICT 541 541
CC FT CONFLICT 622 622
CC FT CONFLICT 622 622
CC FT SEQUENCE 714 AA; 81738 MW; 0BB31DEFC56369A CRC64;
CC -----
Query Match 65.8%; Score 2444.5; DB 1; Length 714;
Best Local Similarity 63.1%; Pred. No. 5,1e-151;
Matches 440; Conservative 120; Mismatches 132; Indels 5; Gaps 3;

```

```

Db 193 KAVAKNGSEYALSGSGTSEGEFDFTGVTWEYLKRAKSDVLSILKALERSLLGCSI 252
Cc 243 DITSADSEALITFOKLVKGAAYSVTGAEEVNSGSLQKILRINPMGEVMTGRANDNC 302
Cc 253 DISVADMEAVITFKLVKGAHAYSVTAKQVNVGGVNLIRNMPGEVMTGAMSDGS 312
Cc 303 SMTTIDPEERERLIRHEDGEFMSFDFLRHYSRLICNLTPDTLTSPTYKKMLTKMD 362
Cc 313 ENMGVDPYORDQLRVAMEDGEFMSFDFLRHYSRLICNLTPDTLTSPTYKKMLTKMD 372
Cc 363 GMRBRSSTAGCGCNVNTWMPQYILKEEDEDDEDE--GSGCTPLVGLIQKRRRO 419
Cc 373 GWRKSTAGGCGCNVNTWMPQYILKEEDEDDEDE--GSGCTPLVGLIQKRRRR 432
Cc 420 RKMGEMHITIGFIVPEPELSCQTHIHLSKNPLFNARERSDTINTREVLNFKLP 479
Cc 433 RFRGRMETITGRAVYVPEPELSCQTHIHLSKNPLFNARERSDTINTREVLNFKLP 491
Cc 480 GEYILVSTPEEPKNDGFCIRVSEKADYQAVDEIEANT--EEFDISDDIDDGVRRLP 538
Cc 492 GEYVVPSTPEEPKEDFVLRFPSEKAGTQELDDQVAILPDEQVLSSEIEDEKALF 551
Cc 539 AQLAGEDAHISAFELQTLIRVLAARQDILKSDFSEITCKINWMLDSGSGTGLTKER 598
Cc 552 ROLAGEDEHISAFELQTLIRVLAARQDILKSDFSEITCKINWMLDSGSGTGLTKER 611
Cc 599 ILWTKIOKTYREIDVDVSGTMSYEMRKLESGFMPQCHQVIVARFADPDLLID 658
Cc 612 ILMNRIRNVLISFRKDLKSGMSAYEMRMALIESGFLKANKKLFILITRISPEPLAVD 671
Cc 659 FDNFVRCVLRLETLFKIFKOLDPENTGTELDLSWL 695
Cc 672 FDNFVRCVLRLETLFKIFKOLDPENTGTELDLSWL 708
Cc -----
RESULT 8
Cc CANI MOUSE STANDARD; PRT; 713 AA.
Cc ID CANI MOUSE STANDARD; PRT; 713 AA.
Cc 16-OCT-2001 (Rel. 40, Created)
Cc 16-OCT-2001 (Rel. 40, Last sequence update)
Cc 28-FEB-2003 (Rel. 41, Last annotation update)
Cc Calpain 1, large (catalytic) subunit (EC 3.4.22.17) (Calcium-activated
Cc neutral proteinase) (CANP) (Mu-type) (mucANP) (micromolar-calpain).
Cc CAPN1 OR CANP1 OR CAPAL.
Cc OS Mus musculus (Mouse)
Cc OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Cc Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Cc NCBI_TaxID=10090;
Cc [1]
Cc SEQUENCE FROM N.A.
Cc RC STRAIN=CF-1;
Cc RA Poirier C., Pousard S., Faust D.M., Imaizumi-Sherer T., Weiss M.C.,
Cc Ducastain A., Montarras D., Pinset C., Guenet J.-L.;
Cc Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
Cc [2]
Cc SEQUENCE FROM N.A.
Cc RC STRAIN=129;
Cc RA Sahr K.E., Andrabi S., Peters L.L., Chishti A.H.;
Cc "Cloning and characterization of the cDNA and gene encoding the mouse
Cc mu-calpain large subunit protein."
Cc Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
Cc -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
Cc catalyze limited proteolysis of substrates involved in
Cc cytoskeletal remodeling and signal transduction.
Cc -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
Cc Arg-Xaa with Leu or Val as the p2 residue.
Cc -1- COFACTOR: Binds 3 calcium ions (By similarity).
Cc -1- ENZYME REGULATION: Activated by micromolar concentrations of
Cc calcium and inhibited by calpastatin.
Cc -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
Cc (regulatory) subunit.
Cc -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma

```

membrane upon Ca++ binding (By similarity).
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF021847; AAB72222.1; -
 DR EMBL: AF084459; AAC33134.1; -
 DR HSSP: Q07009; IDFO.
 DR MEROPS: C02.001; -
 DR MGD; MGI:88263; Capn1.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001300; Protease_C2.
 DR InterPro: IPR000169; SHPOT_acsite.
 DR Pfam: PF01067; Calpain_III; 1.
 DR Pfam: PF00036; ehand; 3.
 DR Pfam: PF00648; Peptidase_C2; 1.
 DR PRINTS: PR00704; CALPAIN.
 DR SMART: SM00720; calpain_III; 1.
 DR SMART: SM00230; Cyepc; 1.
 DR SMART: SM00054; Eph; 3.
 DR PROSITE: PS00018; EF_HAND; 2.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
 FT DOMAIN 30 365 CALPAIN
 FT DOMAIN 366 525 DOMAIN III.
 FT DOMAIN 526 541 LINER.
 FT DOMAIN 542 712 DOMAIN IV.
 FT CA BIND 553 564 EF-HAND 1.
 FT CA BIND 597 608 EF-HAND 2.
 FT CA BIND 627 638 EF-HAND 3.
 FT DOMAIN 662 673 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT DOMAIN 692 703 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT SITE 115 115 BY SIMILARITY.
 FT ACT SITE 272 272 BY SIMILARITY.
 FT ACT SITE 296 296 BY SIMILARITY.
 FT CONFLICT 488 488 L -> P (IN REF. 2).
 FT CONFLICT 496 696 D -> N (IN REF. 2).
 SQ SEQUENCE 713 AA; 82106 MW; 3E1E26C95802B864 CRC64;
 Query Match 65.6%; Score 2435; DB 1; Length 713;
 Best Local Similarity 62.8%; Pred. No. 2.1e-150;
 Matches 437; Conservative 126; Mismatches 129; Indels 4; Gaps 3;

Db 313 ENKVDPYREQARVAKDEGEFPMSPDFRREPTKLEINLPDALKSRILRNMTTFYE 372
 QY 363 GNNRSGTAGGCNNYNTFMNPOYLKLEED- EDEEDSGCFTVLGLOKRROR 420
 Db 373 GTRRSSTAGGCNNYATFVNPQFKLEEVADADYNNREGSCFLLALQKRRER 432
 QY 421 KMGEDHTTGFGIYEPPEELSGQTNTHLSKNFPLTRARERSPTFNLREVLRFPLPG 480
 Db 433 RFGEDMETIGFANYQVPRELAGP-VHLKRDPLANASQAQSHFNLNREVSRIPLPG 491
 QY 481 EYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANL-EEPDISEDIDIDGVRLFA 539
 Db 492 EYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANL-EEPDISEDIDIDGVRLFA 551
 QY 540 QLAGDAEISAFLOTLIRVLAKRQDISDGSITCKIMVMDSDSGKGLKEFY 599
 Db 552 KLAGDMEISVKELOTLIRVLAKRQDISDGSITCKIMVMDSDSGKGLKEFY 611
 QY 600 LMTKIOKQKIVREIVDRSGTNSYEMRKALBEAGFKMPQCHQYVARFADQILDF 659
 Db 612 LMTKIRNYLTIFPKFDIDKSGMSAYEMKALIAAGFKLKLKHLITRYSPLAVDF 671
 QY 660 DNEFRCIVRLTELTKIFKOLDPENTGTIELDLSWL 695
 Db 672 DNEFRCIVRLTELTKIFKOLDPENTGTIELDLSWL 707
 RESULT 9
 CAN1 RAT STANDARD; PRT; 713 AA.
 ID CAN1 RAT
 AC P9571;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated
 neutral proteinase) (CANP) (Mu-type) (mucANP) (Microcatalpain).
 OS CAPN1 OR CLS1.
 OC Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97107433; PubMed=8950173;
 RA Sorimachi H., Amano S., Ishiura S., Suzuki K.;
 RT "Primary sequences of rat mu-calpain large and small subunits are,
 RL respectively, moderately and highly similar to those of human.";
 RL Biochim. Biophys. Acta 1309:37-41(1996).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodelling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions (By similarity).
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding (By similarity).
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U53858; AAC53001.1; -

DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00054; Efb; 3.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; PARTIAL.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; PARTIAL.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; PARTIAL.
KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
FT NON_TER 1 1
FT DOMAIN <1 77 CALPAIN.
FT DOMAIN 78 236 DOMAIN III.
FT DOMAIN 237 251 LINKER.
FT DOMAIN 252 422 DOMAIN IV.
FT CA_BIND 263 274 EF_HAND 1.
FT CA_BIND 307 318 EF_HAND 2.
FT CA_BIND 337 348 EF_HAND 3.
FT DOMAIN 372 383 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT DOMAIN 402 413 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
FT ACT_SITE 8 8 BY SIMILARITY.
SQ SEQUENCE 422 AA; 49494 MW; AE4FA3C48A33C41 CRC64;
Query Match 58.0%; Score 2153; DB 1; Length 422;
Best Local Similarity 93.8%; Pred. No. 2e-132;
Matches 396; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
QY 279 OKLIRINPMGEVMTGRANDNCSPWNTIDPEERERLIRHEDGPFMWSFSDPLRHSRL 338
DB 1 OKLIRINPMGEVMTGRANDNCSPWNTIDPEERERLIRHEDGPFMWSFSDPLRHSRL 60
QY 339 EICNLTPDTLTSPTYKKMKLTKMDGNRRSGTAGCGRNYPNTFMWNPOLYIKLEEDDE 398
DB 61 EICNLTPDTLTSPTYKKMKLTKMDGNRRSGTAGCGRNYPNTFMWNPOLYIKLEEDDEQ 120
QY 399 EDGSGGCTFLVGLQKRRRRKKGEDMTGFGIYEVPELSQNTINHLKSNFFLNTRA 458
DB 121 EDGSGGCTFLVGLQKRRRRKKGEDMTGFGIYEVPELSQNTINHLKSNFFLNTRA 180
QY 459 RERSDTFLNREVNRFPLPGEYILVSTPEPKDGFCFVRSEKKADQAVDEIEA 518
DB 181 RERSDTFLNREVNRFPLPGEYILVSTPEPKDGFCFVRSEKKADQAVDEIEA 240
QY 519 NLEEFDISDDIDDGVRRLFQALAGEPAISAFELQTLIRRLAKRQDIKSGFSIETCK 578
DB 241 NLEEFDISDDIDDGVRRLFQALAGEPAISAFELQTLIRRLAKRQDIKSGFSIETCK 300
QY 579 IMVMMLSDSGGKGLKEFYILMTKIQYQKIYREIDVDRSGTMSYEMRKALBEAGFKM 638
DB 301 IMVMMLSDSGGKGLKEFYILMTKIQYQKIYREIDVDRSGTMSYEMRKALBEAGFKL 360
QY 639 PCQGHQIVARFADDOILIDFNFRCLVRLETFKLPKOLDPENTGTIELDLISWLCFS 698
DB 361 PCQGHQIVARFADDOILIDFNFRCLVRLETFKLPKOLDPENTGTIELDLISWLCFS 420
QY 699 VL 700
DB 421 VL 422

RESULT 11

CANB HUMAN
ID CANB HUMAN STANDARD; PRT; 702 AA.
AC O9UM06;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain 11 (EC 3.4.22.17) (Calcium-activated neutral proteinase 11)
DE (CANP 11).
GN CANP11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9933989; Pubmed=10409436;
RA Dear T.N., Moller A., Boehm T.;
RT "CANP11: A calpain with high mRNA levels in testis and located on
RT chromosome 6";
RL Genomics 59:243-247(1999).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyze limited proteolysis of substrates involved in
CC cytoskeletal remodeling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or
CC Arg|-Xaa with Leu or Val as the P2 residue.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN TESTIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: AJ242832; CAB52137.1; -.
DR HSSP: Q07009; 1DFO.
DR MEROPS; C02.013; -.
DR GENNEW; HGNC:1478; CANP11.
DR MIM; 604822; -.
DR GO; GO:0004198; F:calpain activity; TAS.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease_C2.
DR InterPro; IPR001659; SHPoc_acsite.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyapc; 1.
DR SMART; SM00054; Efb; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
FT DOMAIN 17 352 CALPAIN.
FT DOMAIN 353 514 DOMAIN III.
FT DOMAIN 515 530 LINKER.
FT DOMAIN 531 701 DOMAIN IV.
FT CA_BIND 586 597 EF_HAND 1.
FT CA_BIND 616 627 EF_HAND 2.
FT DOMAIN 651 662 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT ACT_SITE 681 692 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT ACT_SITE 102 102 BY SIMILARITY.
FT ACT_SITE 259 259 BY SIMILARITY.
FT ACT_SITE 283 283 BY SIMILARITY.
SQ SEQUENCE 702 AA; 80582 MW; 3A618682B9674FAE CRC64;
Query Match 54.2%; Score 2014; DB 1; Length 702;
Best Local Similarity 52.4%; Pred. No. 4e-123;
Matches 365; Conservative 133; Mismatches 193; Indels 4; Gaps 2;
QY 4 IAAKLADREAFAGLGSHERAITYLNDYFALNCELEAGTLFQDPSFPAIPALGFKEL 63
DB 1 MVAHINNSRLKAGVGQHDVAQNFQNSFELEPAACLRKELDEDDLPFAEPSSLFKDL 60
QY 64 GPVSSKTRGMRKRPKEICADPOFIIGGARTTICGALGDCWLLAIAISLTNBEITAR 123
DB 61 GPVSSKTRGMRKRPKEICADPOFIIGGARTTICGALGDCWLLAIAISLTNBEITAR 120
QY 124 VVPLNQSFGQNVAGIIFHFGWQYGEWVEVVDRLPTKDGELFLVHSAGSGSFWSALLEK 183

```

Db 121 VVPRGSPKKNVGIHFPOIMFGQWVNVVDRLLPTKNDKLVFVHSTESSEFMSALXK 180
QY 184 AVAKINGCYEALSGATTEGFEDEFTGGIAEWEYELKPPPLPTIOALOKGSLGCSID 243
Db 181 AVAKSGSYALSGGSMELDEFTGGVAVOSFOLQRPQVHLKILRAVRSLSMGCSIE 240
QY 244 ITSAADSEALTPOKLVKGAHAYVTGAEEVNSLSQKILIRINPWEVETGRMNDPCPS 303
Db 241 VTSDESELSWTDKMLVRGHAYVTGLQDVHRYGMETLIRVIRPWGRIENNGAMSDSARE 300
QY 304 WNTIDEBERERLRRRHEDGFPMWSFSDFLNHSRLICNLTPTLTISDYTKMKMLTKMDG 363
Db 301 MEESVADIQWOLHAKHEDGFPMWSYODFLNNFTLLEICNLTPTLTISDYTKMKMLTKMDG 360
QY 364 NMRGSGTAGGCRVYPTFMWNPVOLLLEBEDEDEDESG---CFELVGLIOKHRROR 420
Db 361 SMRSGSAGGCRNHPGFTWNPQKISLPEGDDPDDABGNVAVCTCLVALMOKMNRHAR 420
QY 421 RMGEDMHTIGFGIYEVEPELSGQTNHLSKNFELTNRRERSDPTINLRERLFLKLP 480
Db 421 OQGAQOTIGFVLVAVPKEFQNTQDVHLKKEFTKYQDGHGSEIFLTSRSVSSQLRPPG 480
QY 481 EYILVSTPEPNKDGPCFIVSESEKADYAVD-ELEANLEPDISSEDIDGVRRLPA 539
Db 481 EYIISTEPHNDADFLRVEFKESSEWELDEVNVAEQLOBEKXSEDDMDDFLHLPK 540
QY 540 QLAGEDAEISAFELQTLIRVLAARODIKSDSIETCKIMVMLDSGSKGLKPEFY 599
Db 541 IVAGEKEIGVYVLORLNMAIKFSPKTKGGLDRCMIMLMDKSGKGLLEFKI 600
QY 600 LMTKIOKYOKIYREIDVDRSGTMNSYEMKALEAGFKMPCOLHOVYVAREFADQIIDF 659
Db 601 LMKLKKMMDIFRECQDHSGLTNSYEMRLVIEKAGIKLNNKQVLYVARYADDLIDF 660
QY 660 DNFRVCLVLET.FKIFKQIDPENTGTIELDISWL 695
Db 661 DSFISCFRLKTFWTFEFLTWDPKNTGHCILSLEOWL 696

```

RESULT 12
ID CAN3_CHICK STANDARD; PRT; 810 AA.
AC 092177;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain 3 large subunit (EC 3.4.22.17) (Calpain I3) (Calpain p94,
DE large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
DE (CANP 3) (Muscle-specific calcium-activated neutral proteinase 3 large
DE subunit).
GN CAPN3.
OS Gallus gallus (Chicken).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RT TISSUE=Lung;
RC MEDLINE=95260862; PubMed=7742367;
RA Sorinachi H., Tsukahara T., Okada-Ban M., Sugita H., Ishiura S.,
RA Suzuki K.;
RT Identification of a third ubiquitous calpain species -- chicken
RT muscle expresses four distinct calpains.";
RL Biochim. Biophys. Acta 1261:381-393(1995).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
CC Arg-Xaa with Leu or Val as the P2 residue.
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (potential).
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE. LOW LEVELS IN SPLEEN,
CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D38028; BAA07230.1; -
DR PIR; S57196; S57196.
DR HSRP; Q07009; 1DF0.
DR MEROPS; C02.004; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR000169; SHPoc. acsite.
DR Pfam; PF01067; Calpain III; 1.
DR Pfam; PF00036; ehand_4; 4.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain III; 1.
DR SMART; SM00230; CysPc; 1.
DR SMART; SM00054; Efb; 3.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL PROTEASE CYS; 1.
DR PROSITE; PS00639; THIOL PROTEASE HIS; FALSE NEG.
DR PROSITE; PS00640; THIOL PROTEASE ASN; FALSE NEG.
DR Hydrolase; Thiol protease; Calcium-binding; Nuclear protein;
KW Multi-gene family.
FT DOMAIN 43 421 CALPAIN.
FT DOMAIN 422 579 DOMAIN III.
FT DOMAIN 580 638 LINKER.
FT DOMAIN 639 809 DOMAIN IV.
FT CA_BIND 694 705 EF-HAND 1 (PROBABLE).
FT CA_BIND 724 735 EF-HAND 2 (PROBABLE).
FT DOMAIN 759 770 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT DOMAIN 789 800 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT ACT_SITE 123 123 BY SIMILARITY.
FT ACT_SITE 327 327 BY SIMILARITY.
FT ACT_SITE 351 351 BY SIMILARITY.
SQ
Query Match 810 AA; 93560 MW; E8DE9941IC51041E CRC64;
Best Local Similarity 51.1%; Score 1897; DB 1; Length 810;
Matches 350; Conservative 145; Mismatches 177; Indels 98; Gaps 7;

```

18 LGSHERAIKYNODYALNRNECLEAGTLFQDSEFPAIPSAAGKELGPPSSKTRGRWR 77
Db 41 ISNQPILIKVKEKTYEHLKKEENILVEDDPFPNENSLFYSQRPTR-----PEWR 95
QY 78 PTEICADPOFTIGGARTDICOAGLDCWLLAIASTLTNEETLAVVPLNOSFOENYAG 137
Db 96 PREICENPRFIIGANRTDICOGLDQCFLLAIACTILNKLLICVTHIDGSFIQYAG 155
QY 138 IFHFQWQYGEWVNVVDRLLPTKQDELLFVHSAEGSEFMSALLEKAVAKINGCYEALSG 197
Db 156 IFHFQWQYGEWVNVVDRLLPTKQDELLFVHSAEGSEFMSALLEKAVAKINGCYEALSG 215
QY 198 GATTEGFEDEFTGGIAEWEYELKPPPLPTIOALOKGSLGCSID-----243
Db 216 GNTTEMEFTGGVIVFYELKQAPKDIYIKMKIALARGSLMSIDDNIGFHYGAAPRSO 275
QY 244 -----ITSAADSEALTPOKLVKGAHAYVTGAEEVNSLSQKILIRINPWEVETGRMNDPCPS 270
Db 276 IGLIRARWKNLENAQMTSYVDYOSTDERPAWTIMPQYETRMSCGLVKGHAYSTYAVE 335
QY 271 EYVNSLSQKILIRINPWEVETGRMNDPCPSWNTIDEBERERLRR-HEDEGFPMWSFS 329

QY 258 -----LYKGAYSYTGAEEVESNSLOKLR 283
 Db 296 NSLLRSDSLDPRGSDDRPSTIVVQYETRMACGLVGHAYSYTGLEALFKGEKYLVR 355
 QY 284 IRRPMGEVETGWMNDCPSMNTIDPEERELTRR-HEDEPFMSPSDFLRHYSRLKICN 342
 Db 356 LRRPMQGVENWNGSMDGKMSFVDKDEKARLOHQVTEDEPFMWSYDDFYHFTKLEICN 415
 QY 343 LTPDITLSDTYKKWKLTKMDGNRRGSTAGCCNRYPTFMNPOYLKLEEBDEDEBDE 402
 Db 416 LTPDALLESDELQWTWYSVNEGRRVRCGSAAGCCNRPFTFTNTQYRLKLEEDDDPDS 475
 QY 403 SGGTFLVGLIOKRRRRQKNGEDMTTIGFQIYEVPEELSGQTNILSKNEFLTNRARERS 462
 Db 476 VICSFLVALMQRKRRKRGKLGANLFTIGFAIYEPKEMHGANKQ-HLOKQDFLVNASKARS 534
 QY 463 DTINLRBVLNRKLPGEFILLVSTPEPNKDDFCIRVSEKK-----506
 Db 535 KTYNNKEVSGQRPRLPPESEYIVPSTYEPHQBGFILRVSEKRNLSBAENTISYDRPV 594
 QY 507 -----ADYQAVDEDEANLEB-----FISEDDIDGV 534
 Db 595 KKKKKPITIVSDRANSKELGVDOAEKQKAGPEKGETPOPRPGHTDQSEBQOQF 654
 QY 535 RLPLAQLAGEBAISAFELQTIILRVLAQRQDIKSDGFSIETCKIMVMDLSDSGGGL 594
 Db 655 RNIFRQAGDMEICADELKNVINTVYVNGKDKLQGFITESCRSMALMDTQSGSLNL 714
 QY 595 KEYIILWTQKIQKQYREIDVDRSGTMSYENRKALEEGFKMPCOLHOVIYARFDDQ 654
 Db 715 QEFHMLMKKIKAKQKIFKHVDTGHSCTINSYEMNANVANDGFHINSOLYDITIRYADKH 774
 QY 655 LIIDPFNRCVLRLETLFKIFKQLODPENTGTELDLSML 695
 Db 775 MNIDPDSFICFVRLBGMFRANAFDQDGIITKLNLWML 815

RESULT 14
 ID CAN3_RAT STANDARD; PRT; 821 AA.
 AC P16259;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Calpain 3 large subunit (BC 3.4.22.17) (Calpain I3) (Calpain p94,
 DE large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
 DE (CANP 3) (Muscle-specific calcium-activated neutral protease 3 large
 DE subunit).
 GN CAPN3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=90062125; PubMed=2555341;
 RA Sorimachi H., Imaizoh-Omori S., Emori Y., Kawasaki H., Ohno S.,
 RA Minami Y., Suzuki K.;
 RT Molecular cloning of a novel mammalian calcium-dependent protease
 RT distinct from both m- and mu-types. Specific expression of the mRNA
 RT in skeletal muscle.";
 RL J. Biol. Chem. 264:20106-20111(1989).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Skeletal muscle.

CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC -----
 DR EMBL, J05121; AAA41790.1; -.
 DR F1R, B34488; B34488.
 DR HESP, Q07009; IDPO.
 DR MEROPS, C02_004; -.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001300; Protease_C2.
 DR InterPro: IPR000169; SHPTC_acbSite.
 DR Pfam, PF01067; Calpain_III; 1.
 DR Pfam, PF00036; efhand; 3.
 DR Pfam, PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN-
 DR SMART, SM00720; calpain_III; 1.
 DR SMART, SM00230; Cypec; 1.
 DR SMART, SM00054; Efn; 3.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 DR Hydrolase; Thiol protease; Calcium-binding; Multigene family; Repeat.
 FT DOMAIN 49 428
 FT DOMAIN 429 586
 FT DOMAIN 587 649
 FT DOMAIN 650 820
 FT CA_BIND 705 716
 FT CA_BIND 735 746
 FT DOMAIN 770 781
 FT DOMAIN 800 811
 FT ACT_SITE 129 129
 FT ACT_SITE 334 334
 FT ACT_SITE 358 358
 SQ SEQUENCE 821 AA; 94127 MW; 27FAEAD2FEA19FBF CRC64;
 Query Match 50.4%; Score 1869.5; DB 1; Length 821;
 Best local Similarity 46.1%; Pred. No. 1.2e-113;
 Matches 351; Conservative 131; Mismatches 176; Indels 103; Gaps 8;
 QY 32 YEALRNECLEAGTLFODPSFPALPSALGFKELEGYSKTRGMKRPTEICADPOFTIGG 91
 Db 61 FEQLHKKCLEKVLVLDPEPPEPDETSLFYSQKPEIO-----FWKKRPPEICENPRTIIG 115
 QY 92 ATRTDTCGALGSCWLLAIASTLTNBEILARVPLNOSFOEYVAGIFHFQOYGEWVE 151
 Db 116 ANRTDTCGDLGSCWLLAIACTLTNERLLFRVTPHDQSTFTEYVAGIFHFQRYGDWD 175
 QY 152 VVVDRLPTKDGELLFVHSAEGSEFWSALKEKAYAKINGCYBALSGATTEGEDEFTGI 211
 Db 176 VVDDCLPTNNOLVTKSHRNEFPKALKEKAYALHSGSYELAKKQNTTEAMEDTGGV 225
 QY 212 AEWEYELKPEPNLFKTIQKALQKSLGCSID-----ITSAADEAATTFOK-----257
 Db 236 TEFEIKDAPSDWYKIMRAIERGSLMGCSIDGTMTVYGTSPGLMGLIARVRNMD 295
 QY 258 -----LYKGAYSYTGAEEVESNSLOKLR 283
 Db 296 NSLLRSDSLDPRASDRPSTIVVQYETRMACGLVGHAYSYTGLEALFKGEKYLVR 355
 QY 284 IRRPMGEVETGWMNDCPSMNTIDPEERELTRR-HEDEPFMSPSDFLRHYSRLKICN 342
 Db 356 LRRPMQGVENWNGSMDGKMSFVDKDEKARLOHQVTEDEPFMWSYDDFYHFTKLEICN 415
 QY 343 LTPDITLSDTYKKWKLTKMDGNRRGSTAGCCNRYPTFMNPOYLKLEEBDEDEBDE 402

Db 416 LTADALSDKLTQTVTSVNEGRVWVSGSAGCENFPDTFTWNPVRLKLEEDDDPDSE 475
 Qy 403 SCGTEFLVGLIQKRROROKMGEDMHTTGFYIYEPBELSGOTNHLKSNFLITWRARERS 462
 Db 476 VTCFPLVAMOKNRKDKRLKANLFTTICFALYEPKKEHGNKO-HLQKDFLYNASKARS 534
 Qy 463 DTFINLRVNLRFKLPGEYILVPTFEPNKGDFCIRVFSEKK----- 506
 Db 535 KTYINMEVSGRFRLPPESEYIVPTSTPHQGEFILLVFSEKNLSEANNTISVDRPV 594
 Qy 507 -----ADYQ-----AVDEIFANLEF-----FDISEDDIDDGV 534
 Db 595 KKKKKKPIFVSDRANSKELGVQDEABEGKDKGPKQSGSPORPGHTDQSESEQQOF 654
 Qy 535 RRLPAOLAGEDAEISAFELQTLIRRYLAKRODIDSQFSIETCKIMVMDLSDSGKGL 594
 Db 655 RNTFRQIAGDMELCABDLKAVLNTVYVKKHDKLTQGTSLSCSMILMNTDSSGRNL 714
 Qy 595 KEFYILMTKIQYOKIYREIDVDSGTMNSYEMERKALEBAGFKMPCQLHOVIVARFADQ 654
 Db 715 QEFHLMWKIKAMQKIFPGHYDTDSGTMNSYEMERNAVNDAGFHLNSQLYDITWRADKH 774
 Qy 655 LIIPDNVRLVRLFTFKIFKQIDPENTGTIELDISL 695
 Db 775 MNIDFDFICCFVRLGEMFRAFHAFDKDGDGIKLVLEWL 815

RESULT 15

CAN3 HUMAN STANDARD; PRT; 821 AA.
 AC P20807; Q9YS66; Q9YS67;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 3 large subunit (EC 3.4.22.17) (Calpain I2) (Calpain p94,
 large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
 DE (CAMP 3) (Muscle-specific calcium-activated neutral protease 3 large
 subunit) (ncr-1).
 GN CAPN3 OR CAPN3 OR CAPN3 OR NCCL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND VARIANTS LGMD2A.
 RA Richard I., Broux O., Allmand V., Fougereuse F.,
 RA Chalmatukhichai N., Bourg N., Brenguier L., Devaud C.,
 RA Pasturaud P., Roudaut C., Hillaire D., Passos-Bueno M.-R., Zatz M.,
 RA Tischfield J.A., Fardieu M., Jackson C.E., Cohen D., Beckmann J.S.;
 RT "Mutations in the proteolytic enzyme calpain 3 cause limb-girdle
 RT muscular dystrophy type 2A.";
 RL Cell 81:27-40(1995).
 RN [2]
 RP SEQUENCE OF 44-821 FROM N.A.
 RX MEDLINE=90062125; PubMed=2555341;
 RA Sorimachi H., Imaizumi S., Emori Y., Kawasaki H., Ohno S.,
 RA Mitani Y., Suzuki K.;
 RT "Molecular cloning of a novel mammalian calcium-dependent protease
 RT distinct from both m- and mu-types. Specific expression of the mRNA
 RT in skeletal muscle.";
 RL J. Biol. Chem. 264:20106-20111(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (ISORFMS II AND III).
 RA Dickson J.M.U., Love D., Evans C.W.E.;
 RT "Alternatively exon-spliced isoforms of calpain 3 expressed in human
 RT leukocytes.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP REVIEW ON VARIANTS:
 RX MEDLINE=99264234; PubMed=10330340;
 RA Richard I., Roudaut C., Saenz A., Pogue R., Grimbergen J.E.M.A.,
 RA Anderson L.V.B., Beley C., Cobo A.-M., de Diego C., Eymard B.,

RA Gallano P., Ginjaar H.B., Laea A., Pollitt C., Topaloglu H.,
 RA Urtizberea J.A., de Vliesser M., van der Kooi A., Bushby K., Bakker E.,
 RA Lopez de Munain A., Fardieu M., Beckmann J.S.;
 RT "Calpainopathy -- a survey of mutations and polymorphisms.";
 RL Am. J. Hum. Genet. 64:1524-1540(1999).
 RN [5]
 RP VARIANTS LGMD2A GLN-572 AND GLY-744.
 RX MEDLINE=96176567; PubMed=8624690;
 RA Fardieu M., Hillaire D., Mignard D., Feingold N., Feingold J.,
 RA Mignard D., de Ubeda B., Collin H., Tome F.M.S., Richard I.,
 RA Beckmann J.S.;
 RT "Juvenile limb-girdle muscular dystrophy. Clinical, histopathological
 RT and genetic data from a small community living in the Reunion
 RT Island.";
 RL Brain 119:295-308(1996).
 RN [6]
 RP VARIANTS LGMD2A.
 RX MEDLINE=97294404; PubMed=9150160;
 RA Richard I., Brenguier L., Dincer P., Roudaut C., Bady B.,
 RA Burgunder J.-M., Chemaly R., Garcia C.A., Halaby G., Jackson C.E.,
 RA Kurut D.M., Lefranc G., Legum C., Loiselet J., Merlin L.,
 RA Niveton-Chevallier A., Ollagnon-Roman E., Restagno G., Topaloglu H.,
 RA Beckmann J.S.;
 RT "Multiple independent molecular etiology for limb-girdle muscular
 RT dystrophy type 2A patients from various geographical origins.";
 RL Am. J. Hum. Genet. 60:1128-1138(1997).
 RN [7]
 RP VARIANTS LGMD2A ASN-336; GLN-490; VAL-702 AND GLN-748.
 RX MEDLINE=97410018; PubMed=9266733;
 RA Dincer P., Leturcq F., Richard I., Piccolo F., Yalnizoglu D.,
 RA de Toma C., Akcoeren Z., Broux O., Debbrugrave N., Brenguier L.,
 RA Roudaut C., Urtizberea J.A., Jung D., Tan E., Jeunpierre M.,
 RA Campbell K.P., Kaplan J.-C., Beckmann J.S., Topaloglu H.;
 RT "A biochemical, genetic, and clinical survey of autosomal recessive
 RT limb girdle muscular dystrophies in Turkey.";
 RL Ann. Neurol. 42:222-229(1997).
 RN [8]
 RP VARIANTS LGMD2A ARG-222; GLU-486; TRP-489 AND GLN-748.
 RX MEDLINE=98433925; PubMed=9762961;
 RA Urtasun M., Saenz A., Roudaut C., Poza J.J., Urtizberea J.A.,
 RA Cobo A.-M., Richard I., Garcia Bragado F., Leturcq F., Kaplan J.-C.,
 RA Marti Masso J.F., Beckmann J.S., Lopez de Munain A.;
 RT "Limb-girdle muscular dystrophy in Guipuzcoa (Baeque Country,
 RT Spain).";
 RL Brain 121:1735-1747(1998).
 RN [9]
 RP VARIANT LGMD2A 200-PHE-LEU-204 DEL.
 RX MEDLINE=98112493; PubMed=9452114;
 RA Haeffner K., Speer A., Huebner C., Volt T., Oexle K.;
 RT "A small in-frame deletion within the protease domain of muscle-
 RT specific calpain, p94 causes early-onset limb-girdle muscular
 RT dystrophy 2A.";
 RL Hum. Mutat. Suppl. 1: S298-S300(1998).
 RN [10]
 RP VARIANT LGMD2A GLY-744.
 RX MEDLINE=98317835; PubMed=9655129;
 RA Penisson-Besnier I., Richard I., Dubas F., Beckmann J.S., Fardieu M.;
 RT "Pseudometabolic expression and phenotypic variability of calpain
 RT deficiency in two siblings.";
 RL Muscle Nerve 21:1078-1080(1998).
 RN [11]
 RP VARIANT LGMD2A CYS-360.
 RX MEDLINE=98442618; PubMed=9771675;
 RA Kawai H., Akaike M., Kunishige M., Inui T., Adachi K., Kimura C.,
 RA Kawaijiri M., Nishida Y., Endo I., Kashiwagi S., Nishino H.,
 RA Fujiwara T., Okuno S., Roudaut C., Richard I., Beckmann J.S.,
 RA Miyoshi K., Matsumoto T.;
 RT "Clinical, pathological, and genetic features of limb-girdle muscular
 RT dystrophy type 2A with new calpain 3 gene mutations in seven patients
 RT from three Japanese families.";
 RL Muscle Nerve 21:1493-1501(1998).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or

CC Arg-|-Xaa with Leu or Val as the P2 residue.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=3;
 CC Name=1;
 CC IsoId=P20807-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=P20807-2; Sequence=VSP_005227, VSP_005228;
 CC Name=1;
 CC IsoId=P20807-3; Sequence=VSP_005229;
 CC -1- TISSUE SPECIFICITY: Skeletal muscle.
 CC -1- DISEASE: Defects in CAPN3 are the cause of limb-girdle muscular
 CC dystrophy type 2A (LGMD2A) [MIM:253600]. LGMD2A is both autosomal
 CC dominantly and recessively transmitted. It is characterized by
 CC progressive symmetrical atrophy and weakness of the proximal limb
 CC muscles and elevated serum creatine kinase. The symptoms usually
 CC begin during the first two decades of life, and the disease
 CC gradually worsens, often resulting in loss of walking ability 10
 CC or 20 years after onset.
 CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -1- DATABASE: NAME=Leiden Muscular Dystrophy pages;
 CC NOTE=Calpain-3 mutations in LGMD2A;
 CC WWW=http://www.dmd.nl/capn3_home.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X85030; CAA59403.1; -
 CC EMBL; AF127764; AAD28253.1; -
 CC EMBL; AF127765; AAD28254.3; -
 CC PIR; A56218; C1HDD3.
 CC HSSP; Q07009; IDP0.
 CC MEROPS; C02.004; -
 CC DR GeneW; HGNC:1480; CAPN3.
 CC DR MIM; 114240; -
 CC DR MIM; 253600; -
 CC DR GO; GO:0005622; C:intracellular; TAS.
 CC DR GO; GO:0004198; F:calpain activity; TAS.
 CC DR GO; GO:0008234; F:cysteine-type peptidase activity; TAS.
 CC DR GO; GO:0004871; F:signal transducer activity; TAS.
 CC DR GO; GO:0007517; F:muscle development; TAS.
 CC DR InterPro; IPR002048; EF-hand.
 CC DR InterPro; IPR001300; Protease_C2.
 CC DR InterPro; IPR00169; SHPoc_acsite.
 CC DR Pfam; PF01067; Calpain_III; 1.
 CC DR Pfam; PF0036; Pfam036; 3.
 CC DR Pfam; PF00648; Peptidase_C2; 1.
 CC DR PRINTS; PRO0704; CALPAIN_III; 1.
 CC DR SMART; SM00720; calpain_III; 1.
 CC DR SMART; SM00230; Cyfpc; 1.
 CC DR SMART; SM0054; Efn; 3.
 CC DR PROSITE; PS00018; EF_HAND; 2.
 CC DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 CC DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 CC DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 CC DR HydroLase; Thiol protease; Calcium-binding; Multigene family; Repeat;
 CC Disease mutation; Polymorphism; Alternative splicing.
 CC KW DOMAIN 49 428 CALPAIN.
 CC FT DOMAIN 429 586 DOMAIN III.
 CC FT DOMAIN 587 649 LINKER.
 CC FT DOMAIN 650 820 DOMAIN IV.
 CC FT CA_BIND 705 716 EF-HAND 1 (PROBABLE).
 CC FT CA_BIND 735 746 EF-HAND 2 (PROBABLE).

FT DOMAIN 770 781 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
 FT DOMAIN 800 811 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT ACT SITE 129 129 BY SIMILARITY.
 FT ACT SITE 334 334 BY SIMILARITY.
 FT ACT SITE 358 358 BY SIMILARITY.
 FT VARSPIC 268 315 Missing (in isoform II).
 FT VARSPIC 595 639 /FTId=VSP_005227.
 FT VARSPIC 595 639 /FTId=VSP_005228.
 FT VARSPIC 595 639 /FTId=VSP_005227.
 FT VARSPIC 595 639 /FTId=VSP_005228.

Query Match 50.2%; Score 1862.5; DB 1; Length 821;
 Best Local Similarity 45.5%; Pred. No. 3.3e-113;
 Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;

32 YEALRNECEAGLTPDDSPFPAISALGFELOPYSKTRGMKRTETCAPDQFIIG 91
 61 FEQLHKKCLEKVLVYDPEPDTSLFSQKPPID---FWKKEPPEICENPRFTIDG 115
 92 ATRTDICOGALGDCMLLAIALSLTNEELARVVPINQSPQENYAGIFHPQMGWGE 151
 116 ANRTDICOSELSGCMFLAIACTLTNQHLLFRVPHDQSTIEYAGIFHPQMGWGE 175
 152 VVVDRLPTKDGELFVHSAEGSEFMSALLEKAVANKINGEALSGGATTEGFEPTGI 211
 176 VVIDDCLPTYNQGLVFTKSNHREPMFSALEKAYAKHGSEYALKGANTTEAMEFTGV 235
 212 AEWYELKKPPNLFKTIQKALQGSILGCSID---ITSABSEALTPOK----- 257
 236 AEFPEIRADSDWKIMKKAIERGSLMGCSIDGTMTGTSPSGIMGELLARMVRNMD 295
 258 -----LVGHAAYSTGAEVEESNSLOKILR 283
 296 NSLLQSDLDPRGSDPRPTITIPVQETRMAGLVRGHAISTGLDDEVFKKEXKIVR 355
 284 IRNPGEVETGKWNDCPSWNTIDPEERELTR- HEDGEFMSFSDFLRHSRLICN 342
 356 LRNPQGVENWGSWSRMDKWSFVDKDEKARLOHQTVEGEFMSYEDFIYHFTKLEICN 415
 343 LTPDITLSTPYKKKTKLTKNDGNRRSGTAGGCVNTYPMNQYLKLEEDDEDEGE 402
 416 LTVDALQSDQLQTVTSVNEGRVRCSSAGGCRNFPDTFTWPNQYRLKLEEDDDPDS 475
 403 SGCTPLVGLIQKRRRQKMGEDMTTIGIYVEPELSGQTIHLSKNPFLTRARERS 462
 476 VICSFLVALMQKRRDRKLGASLFTTIGRAIYVPEKMGNKQ-HIQKDFLYNASKARS 534
 463 DTFINLREVLNREKLPGEYIIVSTFEPNKGDFCIRYSEKKADYQAVDEIEANTL-- 520
 535 KTYINRREVSQRRLPSEVIVIPSTYEPHQBSEFLRVFSEKRNISEVENTISYDRPV 594
 521 -----BEFDSEDDIDG----- 533
 595 KKKTKPIIFVSDRANSNKELGVDQES-BEGKGTSPDQKQSPQPOPGSSDQSEBQQQ 653
 534 VRLLPQALGEDAIEAFELQTLRLRVLLAKKODIKSDGSIETCKLMDVLDSDGSGKG 593
 654 FRNIFQIAGDDMEICADELKKYLVVYKHKDLKHGFTLBCSCRMIALMDTDSGSKLN 713
 594 LKEFYILMTKIQYKQIYREIDVRSQTNWSEMRALBEAGKMPCOLHQVTVARFAD 653
 714 LGEFHLNKKIKAMQKIPHYDIDQSGTINSYEMRANVADAGHLNQNQYDITIMVYADK 773
 654 QLIIDPDNFRCLVLETFLKIKQDPENTGTIELDLISWL 695
 774 HNMIDDSFICFVRLEGMFARHARDKQDGIKLVLEWL 815

Search completed: July 24, 2003, 12:52:29
 Job time : 21.3504 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:46:04 ; Search time 34.3469 Seconds
(without alignments)
1959.945 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713

Sequence: 1 MAGIAAKLAKDREARAGLG.....PENTGTIEDLLSMLGFSVL 700

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3705	99.8	700	1 C1H0H2	calpain (EC 3.4.22
2	3518	94.7	700	2 S38361	calpain (EC 3.4.22
3	3098	83.4	700	2 S57194	calpain (EC 3.4.22
4	2579.5	69.5	705	1 C1H0H	calpain (EC 3.4.22
5	2458.5	66.2	714	1 C1H0H	calpain (EC 3.4.22
6	2356.5	63.5	703	2 A48764	calpain (EC 3.4.22
7	2153	58.0	422	2 B24815	calpain (EC 3.4.22
8	1984	53.4	586	2 S57195	calpain (EC 3.4.22
9	1897	51.1	810	1 S57196	calpain (EC 3.4.22
10	1869.5	50.4	821	1 B34488	calpain (EC 3.4.22
11	1862.5	50.2	821	1 C1H0H3	calpain (EC 3.4.22
12	1594.5	42.9	805	1 A55054	calpain (EC 3.4.22
13	1199.5	32.3	758	1 A39343	calpain (EC 3.4.22
14	1184	31.9	653	1 S44749	COG64.2 protein -
15	925.5	24.9	320	2 A24815	calpain (EC 3.4.22
16	839.5	22.6	647	2 UC5772	tissue-specific ca
17	789	21.2	648	2 S71895	sex-determining pr
18	505	13.6	2143	2 G96595	hypothetical prote
19	458	12.3	263	2 A34466	calpain (EC 3.4.22
20	458	12.3	266	1 C1P0L	calpain (EC 3.4.22
21	454	12.2	266	1 C1R0L	calpain (EC 3.4.22
22	450	12.1	268	1 C1H0L	calpain (EC 3.4.22
23	435.5	11.7	184	2 A55143	calpain (EC 3.4.22
24	426	11.5	1597	2 BVF6SL	sol protein, large
25	426	11.5	1597	2 T08428	gene small optic 1
26	394	10.6	634	2 T32871	hypothetical prote
27	380	10.2	616	2 T24825	hypothetical prote
28	291	7.8	805	2 T24431	hypothetical prote
29	278	7.5	617	2 T22175	hypothetical prote

30	274	7.4	623	2 T22177	hypothetical prote
31	240	6.5	198	2 S52094	sorcin - human
32	239	6.4	217	2 A42578	EF-hand calcium-bi
33	236	6.4	198	2 A25706	sorcin - Chinese h
34	208.5	5.6	1134	2 C88040	protein P47F6.5 [l
35	188	5.1	829	2 A82282	hypothetical prote
36	177.5	4.8	425	2 T23122	hypothetical prote
37	171	4.6	611	2 A11941	hypothetical prote
38	168.5	4.5	186	2 D84673	probable calcium b
39	164.5	4.4	153	2 S04970	calcium-binding pr
40	146.5	3.9	335	2 S64352	hypothetical prote
41	146.5	3.9	501	2 A81852	hypothetical prote
42	143	3.9	482	2 S27608	cysteine proteinas
43	134.5	3.6	217	2 T26132	hypothetical prote
44	127.5	3.4	216	2 G70163	hypothetical prote
45	126.5	3.4	169	1 BCKM	calcractin - Chlam

ALIGNMENTS

RESULT 1

C1H0H2
calpain (EC 3.4.22.17) large chain 2 - human
N/Alternate names: calpain chain L-2; calpain II catalytic chain; high-calcium requirin
C/Species: Homo sapiens (man)
C/Date: 21-Nov-1993 #sequence revision 09-Aug-1997 #text_change 16-Jul-1999
C/Accession: S10590; A31218; A33529
R/Sorimachi, H.; Ohm, S.; Emori, Y.; Kawasaki, H.; Saido, T.C.; Ohno, S.; Minami, Y.,
Biol. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990
A/Title: A novel member of the calcium-dependent cysteine protease family.
A/Reference number: S10589; M01D:90380278; PMID:2400579
A/Accession: S10590
A/Molecule type: mRNA
A/Residues: 1-700 <SOR>
R/Imajoh, S.; Aoki, K.; Ohno, S.; Emori, Y.; Kawasaki, H.; Sugihara, H.; Suzuki, K.
Biochemistry 27, 8122-8128, 1988
A/Title: Molecular cloning of the cDNA for the large subunit of the high-Ca(2+)-requiri
A/Reference number: A31218; M01D:8916474; PMID:2852952
A/Accession: A31218
A/Molecule type: protein
A/Residues: 1-210, 'T', 212-394, 'D', 396-445, 'T', 447-700 <IMA>
A/Cross-references: GB:M23254; M01D:9511636; P01D:AAA5645.1; P01D:9511637
A/Note: parts of this sequence were determined by protein sequencing; the amino end of
R/Hata, A.; Ohno, S.; Akita, Y.; Suzuki, K.
U. Biol. Chem. 264, 6404-6411, 1989
A/Title: Tandemly reiterated negative enhancer-like elements regulate transcription of
A/Reference number: A33529; M01D:89197947; PMID:2539381
A/Accession: A33529
A/Molecule type: DNA
A/Residues: 1-67, 'G', 69-72, 'IE', 75-78, 'R', <HAT>
A/Cross-references: DDBJ:J04700; M01D:9179910; P01D:AAA52760.1; P01D:9463086
C/Genetic:
A/Gene: GDB:CAPN2; mCAPN; CAPN1
A/Cross-references: GDB:119750; OMIM:114230
A/Map position: 1pter-1qter
C/Complex: heterodimer of L (large) and S (small) chains
C/Function:
A/Description: catalyzes the hydrolysis of peptides
A/Note: cleaves preferentially after tyrosine, methionine, or arginine residues and bef
C/Specificity: calpain large chain; calmodulin repeat homology; calpain catalytic domai
C/Keywords: acetylated amino end; calcium binding; cysteine proteinase; EF hand; hetero
F:2-700/Product: calpain large chain 2 #status predicted <MAT>
F:75-337/Domains: calpain catalytic domain homology <CLP>
F:529-560/Domains: calmodulin repeat homology <EF1>
F:572-604/Domains: calmodulin repeat homology <EF2>
F:605-634/Domains: calmodulin repeat homology <EF3>
F:637-669/Domains: calmodulin repeat homology <EF4>
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted
F:105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 99.8%; Score 3705; DB 1; Length 700;
Best Local Similarity 99.6%; Pred. No. 3.1e-238;

Matches 697; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAGIAAKIAXDREAEGSHERRAKIKYNDYALNCELEAGTLFODPSPFPAISALGF 60
Db 1 MAGIAAKIAXDREAEGSHERRAKIKYNDYALNCELEAGTLFODPSPFPAISALGF 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOALDGCWLLAASLTNEEI 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOALDGCWLLAASLTNEEI 120
Qy 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTXGDELLFVHSAEGSEFMSAL 180
Db 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTXGDELLFVHSAEGSEFMSAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDFTGGIAEMWELKPPNLFKIIQKALOKSILGC 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGGIAEMWELKPPNLFKIIQKALOKSILGC 240
Qy 241 SIDITSADSEAITFOKLKVGHAYSVTGAEEVNSGLQKILIRINPMGEVEMTGKNDN 300
Db 241 SIDITSADSEAITFOKLKVGHAYSVTGAEEVNSGLQKILIRINPMGEVEMTGKNDN 300
Qy 301 CPSMNTIDPEERERLTRHEDGEFWMSSDFLRHYSRLCINLPTDITSDTYKKMKLT 360
Db 301 CPSMNTIDPEERERLTRHEDGEFWMSSDFLRHYSRLCINLPTDITSDTYKKMKLT 360
Qy 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLKLEEBEDEDEDESGCTFLVGLIQKRRROR 420
Db 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLKLEEBEDEDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMHTIGFGIYEVPEELSGQTNILSKNFLTNRARSDPTINLREVLRPKLP 480
Db 421 KMGEDMHTIGFGIYEVPEELSGQTNILSKNFLTNRARSDPTINLREVLRPKLP 480
Qy 481 EYILVPTFEPPKODFCIRVSEKKADYQAVDDEIANLEEDISEDDIDGVRRLFAQ 540
Db 481 EYILVPTFEPPKODFCIRVSEKKADYQAVDDEIANLEEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIESAFELQTLIRVLAARODIKSDGFSIETCKINWMDLSDSGKLGKEFYIL 600
Db 541 LAGEDAIESAFELQTLIRVLAARODIKSDGFSIETCKINWMDLSDSGKLGKEFYIL 600
Qy 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCQLHQVIVARFADQLIIDF 660
Db 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCQLHQVIVARFADQLIIDF 660
Qy 661 NFVRCLVRLTLFKIFKQIDPENGTIELDLISWLCFSVL 700
Db 661 NFVRCLVRLTLFKIFKQIDPENGTIELDLISWLCFSVL 700

```

RESULT 2

S38361

calpain (EC 3.4.22.17) II large chain - rat

N/Alternate names: calpain II 80k chain

C/Species: Rattus norvegicus (Norway rat)

C/Date: 31-Dec-1993 #sequence revision 02-Aug-1994 #text_change 22-Jun-1999

C/Accession: S38361; S08650; S39751

R/DeLucca, C.I.; Davies, P.L.; Samis, J.A.; Elce, J.S.

Biochim. Biophys. Acta 1216, 81-93, 1993

A/Title: Molecular cloning and bacterial expression of cDNA for rat calpain II 80 kDa subunit

A/Reference number: S38361; MUID:94032492; PMID:8218419

A/Accession: S38361

A/Molecule type: mRNA

A/Residues: 1-700

A/Cross-references: EMBL:109120; NID:g402665; PIDN:AAA6327.1; PID:g402666

R/Samis, J.A.; Back, D.W.; Graham, E.J.; Elce, J.S.

submitted to the EMBL Data Library, February 1990

A/Reference number: S08650

A/Accession: S08650

A/Molecule type: DNA

A/Residues: 380-439 <SAM>

A/Cross-references: EMBL:X51772

C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyc
F/75-327/Domain: calpain catalytic domain homology <CALP>
F/529-560/Domain: calmodulin repeat homology <EF1>
F/572-604/Domain: calmodulin repeat homology <EF2>
F/605-634/Domain: calmodulin repeat homology <EF3>
F/637-669/Domain: calmodulin repeat homology <EF4>
F/105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 94.7%; Score 3518; DB 2; Length 700;

Beat Local Similarity 93.4%; Pred. No. 8.2e-226; Indels 0; Gaps 0;

Matches 654; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

```

Qy 1 MAGIAAKIAXDREAEGSHERRAKIKYNDYALNCELEAGTLFODPSPFPAISALGF 60
Db 1 MAGIAAKIAXDREAEGSHERRAKIKYNDYALNCELEAGTLFODPSPFPAISALGF 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOALDGCWLLAASLTNEEI 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOALDGCWLLAASLTNEEI 120
Qy 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTXGDELLFVHSAEGSEFMSAL 180
Db 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTXGDELLFVHSAEGSEFMSAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDFTGGIAEMWELKPPNLFKIIQKALOKSILGC 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGGIAEMWELKPPNLFKIIQKALOKSILGC 240
Qy 241 SIDITSADSEAITFOKLKVGHAYSVTGAEEVNSGLQKILIRINPMGEVEMTGKNDN 300
Db 241 SIDITSADSEAITFOKLKVGHAYSVTGAEEVNSGLQKILIRINPMGEVEMTGKNDN 300
Qy 301 CPSMNTIDPEERERLTRHEDGEFWMSSDFLRHYSRLCINLPTDITSDTYKKMKLT 360
Db 301 CPSMNTIDPEERERLTRHEDGEFWMSSDFLRHYSRLCINLPTDITSDTYKKMKLT 360
Qy 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLKLEEBEDEDEDESGCTFLVGLIQKRRROR 420
Db 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLKLEEBEDEDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMHTIGFGIYEVPEELSGQTNILSKNFLTNRARSDPTINLREVLRPKLP 480
Db 421 KMGEDMHTIGFGIYEVPEELSGQTNILSKNFLTNRARSDPTINLREVLRPKLP 480
Qy 481 EYILVPTFEPPKODFCIRVSEKKADYQAVDDEIANLEEDISEDDIDGVRRLFAQ 540
Db 481 EYILVPTFEPPKODFCIRVSEKKADYQAVDDEIANLEEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIESAFELQTLIRVLAARODIKSDGFSIETCKINWMDLSDSGKLGKEFYIL 600
Db 541 LAGEDAIESAFELQTLIRVLAARODIKSDGFSIETCKINWMDLSDSGKLGKEFYIL 600
Qy 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCQLHQVIVARFADQLIIDF 660
Db 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCQLHQVIVARFADQLIIDF 660
Qy 661 NFVRCLVRLTLFKIFKQIDPENGTIELDLISWLCFSVL 700
Db 661 NFVRCLVRLTLFKIFKQIDPENGTIELDLISWLCFSVL 700

```

RESULT 3

S57194

calpain (EC 3.4.22.17) large chain 2 - chicken

N/Alternate names: m-calpain heavy chain

C/Species: Gallus gallus (chicken)

C/Date: 28-Oct-1995 #sequence revision 03-Nov-1995 #text_change 20-Jun-2000

C/Accession: S57194

R/Sorimachi, H.; Tsukahara, T.; Okada-Ban, M.; Sugita, H.; Ishiura, S.; Suzuki, K.

Biochim. Biophys. Acta 1261, 381-393, 1995

A/Title: Identification of a third ubiquitous calpain species - chicken muscle expressed

A/Reference number: S57194; MUID:95260862; PMID:7742367

A/Accession: S57194
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-700 <SOR>
 A/Cross-references: EMBL:D38026; NID:9882068; P1DN:BA07228.1; PID:9882069
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: cysteine protease; EF hand; hydrolase
 F/529-560/Domain: calpain catalytic domain homology <CALP>
 F/529-604/Domain: calmodulin repeat homology <EF1>
 F/572-604/Domain: calmodulin repeat homology <EF2>
 F/605-634/Domain: calmodulin repeat homology <EF3>
 F/637-669/Domain: calmodulin repeat homology <EF4>
 F/670-700/Domain: calmodulin repeat homology <EF5>
 F/105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 83.4%; Score 3098; DB 2; Length 700;
 Best Local Similarity 81.1%; Pred. No. 6,5e-198;
 Matches 568; Conservative 67; Mismatches 65; Indels 0; Gaps 0;

```

QY 1 MAGIAAKLAKREAAEGLGSHERAIKYLNODYEALRNCELAGTLFODSPFPAIPSAAGF 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MAGIAAKLAKREAAEGLGSHERAIAAGRHGQAVPYLGQDFGALRECELGRLFTFDPFPAIPSAAGF 60
QY 61 KEIGPVSSKTRGMRKRPTEICADPQITIGATRTDIOGALGDCWLLAIAISLTINEE 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 REIGPMSYKTKGVWCRPTLCSGPRFACGATRTDIOGALGDCWLLAIAISLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFOFMOYGEWVVDDELPTKDELLFVHSAAGSEFMSAL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LARVPLNOSFOENYAGIFHFOFMOYGEWVVDDELPTKNGELLFVHSAAGSEFMSAL 180
QY 181 LERAKVAKINGCYEALSGGATTEGFEFTGGIAEWELKKRPPNLFKIIQKALQKSLG 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 LERAKVAKINGCYEALSGGATTEGFEFTGGIAEWELKKRPPNLFKIIQKALQKSLG 240
QY 241 SIIITSAADSEAITFOQLVKGHAYSVTGAEEVSNGLQKILIRINPGEVEMTGRMN 300
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 SIIITSAADSEAITFOQLVKGHAYSVTGAEEVSNGLQKILIRINPGEVEMTGRMN 300
QY 301 CPSEMTIDPEERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDRLTSDTYKKMKLTK 360
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 CPSEMTIDPEERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDRLTSDTYKKMKLTK 360
QY 361 MDGNWRGRTAGGCRNYPNTFMNPOYLILKEEDEDDEDESGCTFLVGLIOKRRRQR 420
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 MDGNWRGRTAGGCRNYPNTFMNPOYLILKEEDEDDEDESGCTFLVGLIOKRRRQR 420
QY 421 KMGEDMHTTGFALYEVPEPESGQTNILSKNPFILTNARAKSNFTILREVLRFKLPAG 480
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 KMGEDMHTTGFALYEVPEPESGQTNILSKNPFILTNARAKSNFTILREVLRFKLPAG 480
QY 481 EYLVSTPEPNODGFCIRVPESEKKADYQVNDDEIANLEEPDISDDIDDDGVRRLFAQ 540
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 EYLVSTPEPNODGFCIRVPESEKKADYQVNDDEIANLEEPDISDDIDDDGVRRLFAQ 540
QY 541 LAEDDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIWMMDLSDSGSKLGKEFYIL 600
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 LAEDDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIWMMDLSDSGSKLGKEFYIL 600
QY 601 WTKIQYQKQYREIYDVRSSTNYSYEMRRALAEAGFLSCQLHQIIVARPADEDLIDFD 660
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 601 WTKIQYQKQYREIYDVRSSTNYSYEMRRALAEAGFLSCQLHQIIVARPADEDLIDFD 660
QY 661 NPVRCILVRETLFKIFQQLDPENTGTETLILSLCSYVL 700
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 661 NPVRCILVRETLFKIFQQLDPENTGTETLILSLCSYVL 700

```

RESULT 4

CICHH

calpain (EC 3.4.22.17) large chain 4 - chicken

N/Alternate names: calpain catalytic chain; intermediate calcium activated neutral protease; Species: Gallus gallus (chicken)

C/Date: 17-May-1985 #sequence_revision 09-Aug-1997 #extc change 24-Nov-1999
 C/Accession: A00979
 R/Ohno, S.; Emori, Y.; Imajob, S.; Kawasaki, H.; Kitaragi, M.; Suzuki, K.
 Nature 312, 566-570, 1984
 A/Title: Evolutionary origin of a calcium-dependent protease by fusion of genes for a calpain catalytic domain and calmodulin repeat homology
 A/Reference number: A93348; M0ID:85061606; PMID:6095110
 A/Accession: A00979
 A/Molecule type: mRNA
 A/Residues: 1-705 <OHN>
 A/Cross-references: EMBL:X01415; NID:963332; P1DN:CAA25658.1; PID:963333
 R/Emori, Y.; Ohno, S.; Tobita, M.; Suzuki, K.
 FEBS Lett. 194, 249-252, 1986
 A/Reference number: A91354; M0ID:86082358; PMID:3008628
 A/Comments: annotation; gene structure
 C/Content: This calpain has calcium requirements intermediate between those of the high and low molecular weight calpains
 C/Genetics:
 A/Introns: 82/3; 106/1; 145/3; 190/2; 246/3; 274/3; 303/2; 328/2; 382/1; 438/3; 442/3;
 C/Complex: heterodimer of L (large) and S (small) chains
 C/Function:
 A/Description: catalyzes the hydrolysis of peptides
 C/Note: cleaves preferentially after tyrosine, methionine, or arginine residues and before lysine
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: blocked amino end; calcium binding; cysteine protease; duplication; EF hand
 F/78-330/Domain: calpain catalytic domain homology <CALP>
 F/533-564/Domain: calmodulin repeat homology <EF1>
 F/576-608/Domain: calmodulin repeat homology <EF2>
 F/609-638/Domain: calmodulin repeat homology <EF3>
 F/641-673/Domain: calmodulin repeat homology <EF4>
 F/674-705/Domain: calmodulin repeat homology <EF5>
 F/2/Modified site: blocked amino end (Met) (in mature form) #status experimental
 F/108,265,289/Active site: Cys, His, Asn #status predicted

Query Match 69.5%; Score 2579.5; DB 1; Length 705;
 Best Local Similarity 66.6%; Pred. No. 1.8e-163;
 Matches 465; Conservative 111; Mismatches 121; Indels 1; Gaps 1;

```

QY 3 GIAAKLAKREAAEGLGSHERAIKYLNODYEALRNCELAGTLFODSPFPAIPSAAGF 62
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 GIAAKLAKREAAEGLGSHERAIAAGRHGQAVPYLGQDFGALRECELGRLFTFDPFPAIPSAAGF 65
QY 63 LGRYSKTRGMRKRPTEICADPQITIGATRTDIOGALGDCWLLAIAISLTINEE 122
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 LGRYSKTRGMRKRPTEICADPQITIGATRTDIOGALGDCWLLAIAISLTINEE 122
QY 123 RYVPLNOSFOENYAGIFHFOFMOYGEWVVDDELPTKDELLFVHSAAGSEFMSAL 182
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 RYVPLNOSFOENYAGIFHFOFMOYGEWVVDDELPTKDELLFVHSAAGSEFMSAL 182
QY 183 KAVAKINGCYEALSGGATTEGFEFTGGIAEWELKKRPPNLFKIIQKALQKSLG 242
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 KAVAKINGCYEALSGGATTEGFEFTGGIAEWELKKRPPNLFKIIQKALQKSLG 242
QY 243 SIIITSAADSEAITFOQLVKGHAYSVTGAEEVSNGLQKILIRINPGEVEMTGRMN 302
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 SIIITSAADSEAITFOQLVKGHAYSVTGAEEVSNGLQKILIRINPGEVEMTGRMN 302
QY 303 CPSEMTIDPEERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDRLTSDTYKKMKLTK 362
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 CPSEMTIDPEERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDRLTSDTYKKMKLTK 362
QY 363 MDGNWRGRTAGGCRNYPNTFMNPOYLILKEEDEDDEDESGCTFLVGLIOKRRRQR 422
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 MDGNWRGRTAGGCRNYPNTFMNPOYLILKEEDEDDEDESGCTFLVGLIOKRRRQR 422
QY 423 KMGEDMHTTGFALYEVPEPESGQTNILSKNPFILTNARAKSNFTILREVLRFKLPAG 482
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 KMGEDMHTTGFALYEVPEPESGQTNILSKNPFILTNARAKSNFTILREVLRFKLPAG 482
QY 483 EYLVSTPEPNODGFCIRVPESEKKADYQVNDDEIANLEEPDISDDIDDDGVRRLFAQ 541
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 EYLVSTPEPNODGFCIRVPESEKKADYQVNDDEIANLEEPDISDDIDDDGVRRLFAQ 541
QY 541 LAEDDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIWMMDLSDSGSKLGKEFYIL 601
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 541 LAEDDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIWMMDLSDSGSKLGKEFYIL 601

```

Db 546 AGEDEMEISVFELKTLINLVIRARHDKLDKDGSLDCSRMVMNIMDKDGSARLGLVEFOLTM 605

Qy 602 TKIKQKQIYREIVDRSGTMSYEMRKALREAGFKMCCQHQIVYVARFADQILIPDN 661

Db 606 NKIRISWLTIFRQYDLDSGTMSYEMRKALREAGFKMCCQHQIVYVARFADQILIPDN 665

Qy 662 FVRCIVRLLETFLFKIFKQDPDPTGTIELDLISWLCFSV 699

Db 666 FVCCIVRLLETFLFKIFKQDPDPTGTIELDLISWLCFSV 703

RESULT 5

CIRHh

calpain (EC 3.4.22.17) large chain 1 [validated] - human

N:Alternate names: calpain chain L-1; calpain I catalytic chain; low-calcium requiring,

N:Contains: chymotactic factor

C:Species: Homo sapiens (man)

C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text_change 08-Dec-2000

C/Accession: A26213; A36740; S10591

R:Aoki, K.; Imajoh, S.; Ohno, S.; Emori, Y.; Koike, M.; Kosaki, G.; Suzuki, K.

FEBS Lett. 205:313-317, 1986

A/Title: Complete amino acid sequence of the large subunit of the low-Ca²⁺-requiring for

A/Reference number: A26213; MUID:86301172; PMID:3017764

A/Accession: A26213

A/Molecule type: mRNA

A/Residues: 1-714 <AOK>

A/Cross-references: EMBL:X04366; NID:929663; PIDN:CA27881.1; PID:929664

R:Kunitatsu, M.; Higashiyama, S.; Sato, K.; Ohkudo, T.; Sasaki, M.

Biochem. Biophys. Res. Commun. 164, 875-882, 1989

A/Title: Calcium dependent cysteine proteinase is a precursor of a chymotactic factor for

A/Reference number: A36740; MUID:90056492; PMID:2554904

A/Accession: A36740

A/Molecule type: protein

A/Residues: 2-10 <KUN>

A/Experimental source: erythrocytes

R:Sortimachi, H.; Ohmi, S.; Emori, Y.; Kawasaki, H.; Saito, T.C.; Ohno, S.; Minami, Y.; S

Biol. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990

A/Title: A novel member of the calcium-dependent cysteine protease family.

A/Reference number: S10589; MUID:90380278; PMID:2400579

A/Contents: annotation

A/Note: comparison with other gene products

C/Comment: Calpain I is activated by micromolar concentrations of calcium.

C/Genetics:

A/Gene: GDB:CAPN1; mu-CANP

A/Cross-references: GDB:119749; OMIM:114220

A/Map position: 11pter-11qter

C/Complex: heterodimer of L (large) and S (small) chains

C/Function:

A/Description: catalyzes the hydrolysis of peptides

A/Note: cleaves preferentially after tyrosine, methionine, or arginine residues and befo

C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain

C/Keywords: acetylated amino end; calcium binding; cysteine proteinase; duplication; EF

F/2-10/Product: calpain large chain 1 #status predicted <MAT>

F/2-10/Product: chymotactic factor #status experimental <CHIF>

F/85-337/Domain: calpain catalytic domain homology <CALP>

F/542-573/Domain: calmodulin repeat homology <EF1>

F/585-617/Domain: calmodulin repeat homology <EF2>

F/618-667/Domain: calmodulin repeat homology <EF3>

F/650-682/Domain: calmodulin repeat homology <EF4>

F/663-714/Domain: calmodulin repeat homology <EF5>

F/2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

F/115,272,296/Active site: Cys, His, Asn #status predicted

Query Match 66.2%; Score 2458.5; DB 1; Length 714;

Best Local Similarity 63.1%; Pred. No. 2e-155;

Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

Qy 3 GIANKLADREAEAGISGHERAIVLYNDYALNCEAGTLFQDPSPALPALKKE 62

Db 13 GVSAGVQKQARREIGLGHENAIKTLGQDYQLRRCKQSGLTRDEAFPPVPSLGKYD 72

Qy 63 LGPYSSKTRGMWRKPRTEICADPOFIIGATRTDICOQALGDCWLLAIASTLTINEETLA 122

Db 73 LGPNSSKTYGIIKKRPRTEILSNPOFIYDQATRTDICOQALGDCWLLAIASTLTINEETLA 132

Qy 123 RVPPLNOSFENANAGIFHPQFMQYGEVNVYDRLPTDGLLEFPHSGESFNSALLE 182

Db 133 RVPFHQSFPQYAGAGIFHPQFMQYGEVNVYDRLPTDGLLEFPHSGESFNSALLE 192

Qy 183 KAYAKINGCEALSGAGTTEGFTGIAEWELKKPPPNLTKIOLAKOSLLGCSI 242

Db 193 KAYAKINGCEALSGAGTTEGFTGIAEWELKKPPPNLTKIOLAKOSLLGCSI 252

Qy 243 DITSADSEAITFOKLVKGHASVVTGAEEVNSGLQKILIRINPGEVETGRANDNCP 302

Db 253 DISSVLDMEAITFKKLVKGAISVVTGAKQYVNGQVYSLIRMPNPGVEWETGAMSDSS 312

Qy 303 SMTTIDPEERELTRHDEEFPMSPDFLRHRSRLICNLPTDITSPDYKKKXLTMD 362

Db 313 EMMNVDPYERDQLRVKNDEEFPMSPDFLRHRSRLICNLPTDITSPDYKKKXLTMD 372

Qy 363 GNNRSGTAGCCNYPNTFMNPOVYLKLEEDDEBE--DGESCTFLVGLQKHRROR 420

Db 373 GNNRSGTAGCCNYPNTFMNPOVYLKLEEDDEBE--DGESCTFLVGLQKHRROR 432

Qy 421 KMGEDHTTIGFIEVPEELSGQTNILSKNFFLTNRABRSPTFLNREVLNRPYLP 480

Db 433 RFGDMETIGFAYEVPPELVQPAVHLKRDFTLANASRARSBOFTNLREVSRTPLPG 492

Qy 481 EYLLVSTPEPNMDGPPCIRVFESEKADQVADDEIENL--BEFDISEDDIDGVRRLA 539

Db 493 EYVAVSTPEPNMDGPPCIRVFESEKADQVADDEIENL--BEFDISEDDIDGVRRLA 552

Qy 540 QLAGDEAIEAFLOTILRLVLAQKODISDGSIFETCKIMVMDSDSGKGLKEFYI 599

Db 553 QLAGDEAIEAFLOTILRLVLAQKODISDGSIFETCKIMVMDSDSGKGLKEFYI 612

Qy 600 LMTKIOKYIYREIVDRSGTMSYEMRKALREAGFKMCCQHQIVYVARFADQILIPDN 659

Db 613 LMTKIOKYIYREIVDRSGTMSYEMRKALREAGFKMCCQHQIVYVARFADQILIPDN 672

Qy 660 DNFVRCIVRLLETFLFKIFKQDPDPTGTIELDLISWLCFSV 699

Db 673 DNFVRCIVRLLETFLFKIFKQDPDPTGTIELDLISWLCFSV 708

RESULT 6

calpain (EC 3.4.22.17) large chain 2, tissue-specific - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C/Accession: A48764; B48764

R:Sortimachi, H.; Ishiura, S.; Suzuki, K.

J. Biol. Chem. 268, 19476-19482, 1993

A/Title: A novel tissue-specific calpain species expressed predominantly in the stomach

A/Reference number: A48764; MUID:93374936; PMID:7690035

A/Accession: A48764

A/Status: preliminary

A/Molecule type: DNA; protein

A/Residues: 1-703 <SOR>

A/Cross-references: GB:014479; NID:9441199; PIDN:BA040370.1; PID:9441200

A/Experimental source: stomach

A/Note: sequence extracted from NCBI backbone (NCBIN:137770, NCBIP:137771)

A/Accession: B48764

A/Status: preliminary

A/Molecule type: DNA; protein

A/Residues: 1-379, 'SS' <SO>

A/Cross-references: GB:014480; NID:9441201; PIDN:BA040371.1; PID:94495223

A/Experimental source: stomach

A/Note: sequence extracted from NCBI backbone (NCBIN:137773, NCBIP:137775)

C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain

C/Keywords: alternative splicing; cysteine proteinase; EF hand; hydrolase

F/75-327/Domain: calpain catalytic domain homology <CALP>

F/532-563/Domain: calmodulin repeat homology <EF1>

F/575-607/Domain: calmodulin repeat homology <EF2>

F/640-672/Domain: calmodulin repeat homology <EF3>
F/105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 63.5%; Score 2356.5; DB 2; Length 703;
Best Local Similarity 60.9%; Pred. No. 1,2e-148;
Matches 427; Conservative 127; Mismatches 140; Indels 7; Gaps 3;

```

QY 1 MAGIAAKLAKREAEAGLSHERAIXYLNODYALRNECEAGTLFODPSPAPALPSALGF 60
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 MALAAGVSKORAVAEGLSGNNAVKYLGODFETLRKQCLNSGVLFPDPEFPACPSALGY 60
QY 61 KEIGPSSKTRGRMRKRPTEICADPPITIGATRTDICGALGDCWLLAIASTLTNEE1 120
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 61 KDLGPSPPTQGIYWKAPTELCPNPQPIVGAATRTDIRQGLDPCWLLAIASTLTNEKL 120
QY 121 LAEVLPLNOSFOENYAGIFHFQFMOYGEWEVVDRLPTKDGELFVHSAEGSEFMSAL 180
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 121 LRVVLPRDGSFOQDIYHFPQWQYGEWEVVIDRLPTKQGLFLHSEBEGNEFMSAL 180
QY 181 LEKAVAKINGCYALSGGATTEGFEPTGIAEMWELKPPNLFKIIQKALQKSGSLGC 240
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 181 LEKAVAKINGCYALSGGATTEGFEPTGIAEMWELKPPNLFKIIQKALQKSGSLGC 240
QY 241 SIDITSAADSEATTPQKLVKGAHYSTGAEVSNGLQKLIIRNPWGEVETGRMND 300
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 241 SIDVSTAEEAETTRQKLVKGAHYSTGAEVSNGLQKLIIRNPWGEVETGRMND 300
QY 301 CPENMTIDPERERLRLRRHEDGEFMMSPDELRYHSLEICNLTPDTLSDTYKKMKLTK 360
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 301 APENMTIDPERERLRLRRHEDGEFMMSPDELRYHSLEICNLTPDTLSDTYKKMKLTK 360
QY 361 MDGNMRGRTAGGCRNYPNTFMNPOYLKLEBEDDEDEDEG--ESGCTPLVGLIQRKRR 418
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 361 FNGRMTRGRTAGGCRNYPNTFMNPOYLKLEBEDDEDEDEG--ESGCTPLVGLIQRKRR 418
QY 419 QRMKGEDMTIGGIYVEPELSGQTNILSKNPFELTNRERSDTFINREVLNRPKLP 478
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 419 QRMKGEDMTIGGIYVEPELSGQTNILSKNPFELTNRERSDTFINREVLNRPKLP 478
QY 421 QKRIGQMLSIGYAVQIPEKLESHDPAHLGRDFLGRDPTCSSTYMNIREVSRVRLP 480
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 421 QKRIGQMLSIGYAVQIPEKLESHDPAHLGRDFLGRDPTCSSTYMNIREVSRVRLP 480
QY 479 PGELYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEANLE--FDISEDDIDQVR 535
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 479 PGELYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEANLE--FDISEDDIDQVR 535
QY 481 PGELYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEANLE--FDISEDDIDQVR 538
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 481 PGELYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEANLE--FDISEDDIDQVR 538
QY 536 RLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCKIMVMDLSDSGSLGKLG 595
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 536 RLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCKIMVMDLSDSGSLGKLG 595
QY 539 SLPEEFGVQKSEISANQKVLNEVLSKRTDMKFDGNITCEMISLSDSGSLGEM 598
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 539 SLPEEFGVQKSEISANQKVLNEVLSKRTDMKFDGNITCEMISLSDSGSLGEM 598
QY 596 EPIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKMPCQLHQIVARFADQOL 655
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 596 EPIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKMPCQLHQIVARFADQOL 655
QY 599 EFKTLMLKIRTYLEIFQEMDMHNHVGITAEHEMTALKKAGFTLNNQVQTIARVACSL 658
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 599 EFKTLMLKIRTYLEIFQEMDMHNHVGITAEHEMTALKKAGFTLNNQVQTIARVACSL 658
QY 656 IIDFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLC 696
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 656 IIDFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLC 696
QY 659 GVDNFGVACMIRLETLFKFLILDKQNGIYOLSLAEWLC 699
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 659 GVDNFGVACMIRLETLFKFLILDKQNGIYOLSLAEWLC 699

```

RESULT 7

B24815
calpain (BC 3.4.22.17) large chain 2 - rabbit (fragment)
N/Alternate names: calcium-activated neutral proteinase (CANP)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #text_change 16-Jul-1999
C/Accession: B24815
R/Author: Y.; Kawasaki, H.; Sugihara, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.
J. Biol. Chem. 261, 9465-9471, 1996
A/Title: Isolation and sequence analyses of cDNA clones for the large subunits of two 18
A/Reference number: A92594; MUID:86250902; PMID:2424911
A/Accession: B24815
A/Molecule type: mRNA
A/Residues: 1-422 <EMO>
A/Cross-references: GB:M13797; NID:g165665; PIDN:AAA1455.1; PID:g165666
C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: calcium binding; cysteine proteinase; EF hand; hydrolase

F/251-282/Domain: calmodulin repeat homology <EF1>
F/294-326/Domain: calmodulin repeat homology <EF2>
F/327-356/Domain: calmodulin repeat homology <EF3>
F/359-391/Domain: calmodulin repeat homology <EF4>

Query Match 58.0%; Score 2153; DB 2; Length 422;
Best Local Similarity 93.8%; Pred. No. 1,9e-135;
Matches 396; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

```

QY 279 QKLIIRNPWGEVETGRMNDNCPNMTIDPERERLRLRRHEDGEFMMSPDELRYHSRL 338
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 QKLIIRNPWGEVETGRMNDNCPNMTIDPERERLRLRRHEDGEFMMSPDELRYHSRL 338
QY 339 EICNLTPDTLSDTYKKMKLTKMDGNMRGRTAGGCRNYPNTFMNPOYLKLEBEDDE 398
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 339 EICNLTPDTLSDTYKKMKLTKMDGNMRGRTAGGCRNYPNTFMNPOYLKLEBEDDE 398
QY 61 EICNLTPDTLSDTYKKMKLTKMDGNMRGRTAGGCRNYPNTFMNPOYLKLEBEDDEQ 120
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 61 EICNLTPDTLSDTYKKMKLTKMDGNMRGRTAGGCRNYPNTFMNPOYLKLEBEDDEQ 120
QY 399 EDGESCTPLVGLIQRKRRQRKMGEDMTIGGIYVEPELSGQTNILSKNPFELTNRA 458
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 399 EDGESCTPLVGLIQRKRRQRKMGEDMTIGGIYVEPELSGQTNILSKNPFELTNRA 458
QY 459 RERSDPTFINREVLNRPKLPGEYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEA 518
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 459 RERSDPTFINREVLNRPKLPGEYIIVPSTFEENKGDPCIRVSEKKADYQAVNDEIEA 518
QY 519 NLEEFIDSEDDIDQVRRFLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCK 578
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 519 NLEEFIDSEDDIDQVRRFLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCK 578
QY 241 DLEADVSEDDIDQVRRFLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCK 300
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 241 DLEADVSEDDIDQVRRFLFAQAGDAEISAELOTLIRVLAKRODICKSGFSIETCK 300
QY 579 IMVMDLSDSGSLGLKEFIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKM 638
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 579 IMVMDLSDSGSLGLKEFIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKM 638
QY 301 IMVMDLSDSGSLGLKEFIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKM 360
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 301 IMVMDLSDSGSLGLKEFIYLTMTKIQKYIREIDVDRSGTMSYEMKALEEAGFKM 360
QY 639 PCQLHQIVARFADQOLIIFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLCFS 698
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 639 PCQLHQIVARFADQOLIIFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLCFS 698
QY 361 PCQLHQIVARFADQOLIIFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLCFS 420
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 361 PCQLHQIVARFADQOLIIFQNPVRLVLETLFKFKOLDPENTGTFIELDLSWLCFS 420
QY 699 VL 700
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 699 VL 700
QY 421 VL 422
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 421 VL 422

```

RESULT 8

S57195
calpain (BC 3.4.22.17) large chain 1 - chicken (fragment)
N/Alternate names: mu-calpain heavy chain
C/Species: Gallus gallus (chicken)
C/Date: 28-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C/Accession: S57195
R/Author: H.; Tsukahara, T.; Okada-Ban, M.; Sugita, H.; Iehura, S.; Suzuki, K.
Biochim. Biophys. Acta 1261, 381-393, 1995
A/Title: Identification of a third ubiquitous calpain species - chicken muscle expressed
A/Reference number: S57194; MUID:95260862; PMID:7742367
A/Accession: S57195
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-586 <SOR>
A/Cross-references: EMBL:D38027; NID:g882070; PID:g882071
C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: cysteine proteinase; EF hand; hydrolase
F/1-208/Domain: calpain catalytic domain homology (fragment) <CALP>
F/457-489/Domain: calmodulin repeat homology <EF1>

Query Match 53.4%; Score 1984; DB 2; Length 586;
Best Local Similarity 61.0%; Pred. No. 5,1e-124;
Matches 354; Conservative 108; Mismatches 114; Indels 4; Gaps 2;

```

QY 120 ILARVPLNOSFOENYAGIFHFQFMOYGEWEVVDRLPTKDGELFVHSAEGSEFMSA 179
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 ILARVPLNOSFOENYAGIFHFQFMOYGEWEVVDRLPTKDGELFVHSAEGSEFMSA 179
QY 180 LEKAVAKINGCYALSGGATTEGFEPTGIAEMWELKPPNLFKIIQKALQKSGSLGC 239
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 180 LEKAVAKINGCYALSGGATTEGFEPTGIAEMWELKPPNLFKIIQKALQKSGSLGC 239

```


Db 61 LLEKAVKAVNGCYEALSGSTSEGFEDFTGVTWYDLRKRPADLYOINIKALERSGLG 120

Qy 240 CSIDITSAADSEAITFOKLYKHAAYVTGAEEVENSGSLQKILIRNPMGWVETGMRND 299

Db 121 CSIDITSAADSEAITFOKLYKHAAYVTGAEEVENSGSLQKILIRNPMGWVETGMRND 180

Qy 300 NCPSWNTIDPEERELTRRHEDGEFWMSSFDPLRHYSRLICNLTPDTLTSDTYKWKLT 359

Db 181 SSSQWNEVESRLQOINWMEDEGEFWMSSFDPLRHYSRLICNLTPDTLTSDTYKWKLT 240

Qy 360 KMDGNMRGSGTAGCCRRYPTFMNPOYLKLEEDDEDEP--GEGGCTFLVLQIKR 416

Db 241 LYDSWMRGSGTAGCCRRYPTFMNPOYLKLEEDDEDEP--GEGGCTFLVLQIKR 300

Qy 417 RROKMGEDMHTTIFGVEVEEELSGQTNHLSKNFELTNARERSPTFLIRVLRPFK 476

Db 301 RRRRYGKDMETTFYAYVEVEPEHVGSGVHLQDFELSNARSRSEOFINLRVSTRLR 360

Qy 477 LPPGEYILVPSFEPNKGDFCIRVFSEKADYQAVDEIEANL--EEFDISEDDIDDGVR 535

Db 361 LPPGEYILVPSFEPNKGDFCIRVFSEKADYQAVDEIEANL--EEFDISEDDIDDGVR 420

Qy 536 RLPLQAGEADAEISAFELQTLIRVLAQRDIXSDGSIETCKIMVMDLSDSGKGLK 595

Db 421 ALFROLAGPDMELSVTELOTILNRIIAKHDLRTKGFSLBSCRSVMIMLMDGKGLGLV 480

Qy 596 EFTYLMTRIKOKYKIREIDVRSGTMSYEMRKALEAGFMPCOLHOVTVARFADOL 655

Db 481 EFTYLMTRIKOKYKIREIDVRSGTMSYEMRKALEAGFMPCOLHOVTVARFADOL 540

Qy 656 IIDPNEFVRCVLRLETLFKIFKQIDPENTGTIELDLISWL 695

Db 541 AIDFDSFVCCVLRLETLFKIFKQIDPENTGTIELDLISWL 580

RESULT 9

SS7196
calpain (EC 3.4.22.17) large chain 3 - chicken
N/Alternate names: calpain p94 heavy chain; n-calpain-1 heavy chain
C/Species: Gallus gallus (chicken)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: SS7196
R/Sorimachi, H.; Tsukihara, T.; Okada-Ban, M.; Sugita, H.; Ishiura, S.; Suzuki, K.
Biochim. Biophys. Acta 1261, 381-393, 1995
A/Title: Identification of a third ubiquitous calpain species - chicken muscle expressed
A/Reference number: SS7194; MUID:95260862; PMID:7742367
A/Accession: SS7196
A/Molecule type: mRNA
A/Status: preliminary
A/Residues: 1-810 <SOR>
A/Cross-references: EMBL:D38028; NID:G882072; PIDN:BA07230.1; PID:G1552167
C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: cysteine protease; EF hand; hydrolase
F/93-333/Domain: calpain catalytic domain homology <CALP>
F/746-778/Domain: calmodulin repeat homology <EFH>

Query Match 51.1%; Score 1897; DB 1; Length 810;
Best Local Similarity 45.5%; Pred. No. 4.8e-118;

Matches 350; Conservative 145; Mismatches 177; Indels 98; Gaps 7;

Qy 18 LGSHERAIKYLNDQYEALENECEAGTLFQDPSPFAIPALGFYELGYSSKTRGMWR 77

Db 41 ISRNQPIIKVKEKYEELHKKCLLENILYEDPDPNPTSIFYQKVPK-----FEK 95

Qy 78 PTEICADPQITIGATRTDTCQALGDCWLLAIAISLTNEELIARVPLNQSFOENYAG 137

Db 96 PREICEPRFIIGANRTDTCQALGDCWFLAIAIACTLNLKCLCRVPHQOSFIONYAG 155

Qy 138 IFHQFOYQGEWVAVVDDRLPTKQGLLFVHABSGFPMALKEKVAKINGCYEALSG 197

Db 156 IFHQFOYQGEWVAVVDDRLPTKQGLLFVHABSGFPMALKEKVAKINGCYEALSG 215

Qy 198 GATTEGEDFTGIAEWELKKPPNLFKIQLAKQKSLGCSID----- 243

Db 216 GNTTEAMEDFTGVIIFYEIKDAPKDIYKIMKAIARGSLMSSIDNLFHYGAAPRS 275

Qy 244 -----LISAADSEAITFOK-----LYKHAAYVTGA 270

Db 276 IGEILIRNPMGWVETGMRNDNCPSWNTIDPEERELTRR--HEDGEFWMSS 335

Qy 271 EVESNGSLQKILIRNPMGWVETGMRNDNCPSWNTIDPEERELTRR--HEDGEFWMSS 329

Db 336 ETTKGEKORLVRNPMGWVETGMRNDNCPSWNTIDPEERELTRR--HEDGEFWMSS 395

Qy 330 DPLRHYSRLICNLTPDTLTSDTYKWKLTMDGNMRGSGTAGCCRRYPTFMNPOYL 389

Db 396 DPLRHYSRLICNLTPDTLTSDTYKWKLTMDGNMRGSGTAGCCRRYPTFMNPOYL 455

Qy 390 KLEEDDEDEDESGGCTFLVLQIKRRLRORRORRORRORRORRORRORRORRORR 449

Db 456 KLEEDDEDEDESGGCTFLVLQIKRRLRORRORRORRORRORRORRORRORRORR 514

Qy 450 KNEFLTNARERSDFTINLRVLRPFKLPGEYILVPSFEPNKGDFCIRVFSEKADY 509

Db 515 KDFFLYNASAKRSKTYINNRERISERFLPSEVVIIPSTYEPHQEGEFLIRVSEKRS 574

Qy 510 QAVDEIEANL-----EEFDISEDDIDDG----- 533

Db 575 EEVENMIADRPSEKKGKPIIFVSDRANSNKELTTDEBAGKGEKTHVDEKRS 634

Qy 534 -----VRLFAOLAGEADAEISAFELQTLIRVLAQRDIXSDGSIETCKIMVMDL 585

Db 635 EKSEETOPFNIRQLAGDMEICRELRVNLNVKXKDLTBGELBSSRSMALMD 694

Qy 586 SDGSGKGLKEFYLTWKIOKYKIREIDVRSGTMSYEMRKALEAGFMPCOLHOV 645

Db 695 TDSGKINPDEFHMDKIKSMQKIFKHYDADHSGTINSTEMNNAVKDAGFRNLNDLY 754

Qy 646 IVARFADOLIIDPNEFVRCVLRLETLFKIFKQIDPENTGTIELDLISWL 695

Db 755 ITWRVADKMNNDIDFDFICFVRLDAMFRAHFDXGDGIILKLVLEWL 804

RESULT 10

B34488
calpain (EC 3.4.22.17) large chain 3 - rat
N/Alternate names: Rattus norvegicus (Norway rat)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: B34488; S10589
R/Sorimachi, H.; Imajo-Ohts, S.; Emori, Y.; Kawasaki, H.; Ohno, S.; Minami, Y.; Suzuki, J.
Biochem. Biophys. Res. Commun. 264, 20106-20111, 1999
A/Title: Molecular cloning of a novel mammalian calcium-dependent protease distinct from
A/Reference number: A94688; MUID:90062125; PMID:2555341
A/Accession: B34488
A/Molecule type: mRNA
A/Residues: 1-821 <SOR>
A/Cross-references: GB:005121; NID:G205955; PIDN:AAA1790.1; PID:G205956
R/Sorimachi, H.; Ohmi, S.; Emori, Y.; Kawasaki, H.; Saido, T.C.; Ohno, S.; Minami, Y.; Suzuki, J.
Biol. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990
A/Title: A novel member of the calcium-dependent cysteine protease family.
A/Reference number: S10589; MUID:90380278; PMID:2400579
A/Accession: S10589
A/Status: preliminary
A/Molecule type: mRNA
C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: calcium binding; cysteine protease; EF hand; hydrolase
F/99-400/Domain: calpain catalytic domain homology <EF1>
F/649-680/Domain: calmodulin repeat homology <EF2>
F/692-724/Domain: calmodulin repeat homology <EF3>
F/725-754/Domain: calmodulin repeat homology <EF4>
F/757-789/Domain: calmodulin repeat homology <EF5>
F/790-821/Domain: calmodulin repeat homology <EF6>
F/129,334,358/Active site: Cys, His, Asn #status predicted

Db 654 FRNFKOIAAGDMIEICADELKKVLTNVNKKHDKLTHGFTLES CRSMIATMDTGGSKLN 713
 QY 594 LKEEYLIWTKIQOKQIYREIDVDKSGTMSYEMRKALKEAGFQPCOLHOVIYARPAD 653
 Db 714 LQEHHLNNKIKAMQKFKHYDIDSGTINSYENRNANVNDGFHANNOLYDIITRYADK 773
 QY 654 QLIIDFNPVACLVRLETFKIFKQIDPENTGTELDLISWL 695
 Db 774 HNNIDFDSFICCFVRLBGMFRAFIAPFDGDIITKLVNLEWL 815

RESULT 12

A55054
 calpain (EC 3.4.22.17) large chain - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
 C/Accession: A55054
 R/Emori, Y.; Saito, K.
 J. Biol. Chem. 269, 25137-25142, 1994
 A/Title: Calpain localization changes in coordination with actin-related cytoskeletal ch
 A/Reference number: A55054; MUID:95014293; PMID:7929201
 A/Accession: A55054
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-805 <EMO>
 A/Cross-references: GB:X78555; NID:g562287; PIDN:CAAS5297.1; PID:g562288
 C/Genetics:
 A/Gene: FlyBase:Calpa
 A/Cross-references: FlyBase:FBgn0012051
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: cysteine proteinase; duplication; EF hand; hydrolyase
 F/90-347/Domain: calpain catalytic domain homology <CALP>
 F/676-708/Domain: calmodulin repeat homology <EF1>
 F/709-738/Domain: calmodulin repeat homology <EF2>
 F/741-773/Domain: calmodulin repeat homology <EF3>
 F/774-805/Domain: calmodulin repeat homology <EF4>

Query Match 42.9%; Score 1594.5; DB 1; Length 805;
 Best Local Similarity 41.2%; Pred. No. 6e-98;
 Matches 322; Conservative 133; Mismatches 209; Indels 117; Gaps 14;

QY 13 EAAELGSHERAIVKYNODYALNNECLAEGLTFQDPSFPAIPALGFKELPYSSKTRG 72
 Db 38 EKSSSLRPPYSEV-----QDYETIINSCIASGLFEDPLFPASNESLQGRS-----RPRH 87
 QY 73 MRWRPREICADPQPIIGGATRTDICGALGDCWLLAASLTNBEILAAVPLNLSQ 132
 Db 88 IEWLRPHIAENPQFVEGYSRFPVQOGLDCWLLATATNLQESNLEFFVITPAEQSFE 147
 QY 133 ENYAGIFHFQFQWGEWVVDRLPTKQDELLFVHSAEGSEFMSALLEKAVAKINGCY 192
 Db 148 ENYAGIHFHRFQWQKWDVVIDRLPTYNGLMWMSTENKFRMSALLEKAVAKINGCY 207
 QY 193 EALSAGATGEGEDFTGIAEWELKRPPLFKIIQKALOKSGILGCSIDITSAADSEA 252
 Db 208 EALGSGSTCEMEBFTGVSSEWYDLKEAPGNLFTILQKAARNMGCISIE-PDPNTEA 266
 QY 253 ITPQKLVKGAHYSTTGAEVE-----SNGLQKILIRINPVG-EVETWGRNNDCCPMNT 306
 Db 267 ETPQGLRGHAYSTTKCLDITVTPNRQKI-PMIRMNPMGNAENWNGVWSDSSPERY 325
 QY 307 IDPERERLTRRH-DGEFWMFSFDPLRHYRLIEICNLTPTLTSDTY----KKMKLTLM 361
 Db 326 IPEQKXAEIGLTPRDEGFWMFSFODFLNHPRVEICNLSPTLTEDQHSKRRKEMSMY 385
 QY 362 DGNRRSGTAGGCNYPVTFMNPQYILKEEEDDEDESGCTFLVGLIQQHRRQRK 421
 Db 386 EGEWTPVATAGGCNPLDTPFHNPOYIITLVDPDEDEEGQ--CTVIAALMOKRRSRKN 443
 QY 422 MGEEMHTIGRIYVPE-ELSGQINIHLSKNFPLTNARESDPFINIREVLNFKLPBG 480
 Db 444 MGECLTITGFIYSLNDRLENRPO--GLNFFRYKSSVGRSPHINTREVCAKFKLPBG 500

QY 481 EYILVSTPEPNKDGFCIRVSEKKADYQAVDEIEANLEEFISE-----DDIDD 532
 Db 501 HYILVSTPDPNEGEFIRVFSSETQNMNEENDHVGAGKADTIRPGFPIDPQKE 560
 QY 533 GVRRLPAQLAGEPAEISABELQTLIRVLAKODI----- 567
 Db 561 GLRRLPDSIAGKMEVDMMELKRLDH--SMRDLKRPVFNFRFSNNMAFETQAAGPDD 618
 QY 568 -----KSDGFSI 574
 Db 619 GAGACGLISLIGCPFLKGTPEEQQLGNDQSNKRLIGDNPADGGPYTANAIYDETHGFSK 678
 QY 575 ETCKINVDMDSDSGSKLGKEFYLIWTKIQOKQIYREIDVDKSGTMSYEMRKALKEA 634
 Db 679 DVCRSWVAMLDADKSGKLGKEFEPTLSEIAKKALFKYVDVENTGRVSGFQRLBALNSA 738
 QY 635 GFKMPQOLHOVIYARPADQOLIIDFNPVACLVRLETFKIFKQIDPENTGTELDLISW 694
 Db 739 GYHNNRVLNVLGHRRGSRDGKIAFDFTMCVAKIKTYIDIFKERTKNETATFTLEW 798
 QY 695 L 695
 Db 799 I 799

RESULT 13

A39343
 calpain (EC 3.4.22.17) large chain - fluke (Schistosoma mansoni)
 C/Species: Schistosoma mansoni
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A39343; A45642
 R/Andersen, K.; Tom, T.D.; Strand, M.
 J. Biol. Chem. 266, 15085-15090, 1991
 A/Title: Characterization of cDNA clones encoding a novel calcium-activated neutral prot
 A/Reference number: A39343; MUID:91332027; PMID:1669543
 A/Accession: A39343
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-758 <AND>
 A/Cross-references: GB:M67499; NID:g160936; PIDN:AAA29858.1; PID:g160937
 R/Karcz, S.R.; Podesta, R.B.; Siddiqui, A.A.; Dekaban, G.A.; Strejan, G.H.; Clarke, M.W
 Mol. Biochem. Parasitol. 49, 333-336, 1991
 A/Title: Molecular cloning and sequence analysis of a calcium-activated neutral protease
 A/Reference number: A45642; MUID:92131071; PMID:1775175
 A/Accession: A45642
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-14, 'A', 16-120, 'A', 121-127, 'H', 129-213, 'R', 215-384, 'VTC', 388-440, 'S', 442-758
 A/Cross-references: GB:M74233; NID:g160934
 A/Note: sequence extracted from NCBI backbone (NCBI:79194)
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: calcium binding; cysteine proteinase; duplication; EF hand; hydrolyase
 F/14-380/Domain: calpain catalytic domain homology <CALP>
 F/628-660/Domain: calmodulin repeat homology <EF1>
 F/661-690/Domain: calmodulin repeat homology <EF2>
 F/693-726/Domain: calmodulin repeat homology <EF3>
 F/727-758/Domain: calmodulin repeat homology <EF4>
 F/14,313,337/Active site: Cys, His, Asn #status predicted

Query Match 32.3%; Score 1199.5; DB 1; Length 758;
 Best Local Similarity 38.6%; Pred. No. 9.5e-72;
 Matches 279; Conservative 136; Mismatches 234; Indels 73; Gaps 26;

QY 16 EGLGSHERAIVKYNODYALNNECLAEGLTFQDPSFPAIPALG-----FKELOPYSSKTR 71
 Db 70 KQAKFLMNVNNAVKQYETLVKRLKTERLTWEDPDPADKALGNLPDFE----- 120
 QY 72 GMEKRPTEICADPQPIIGGATRTDICGALGDCWLLAASLTNBEILAAVPLNLSQ 131
 Db 121 RIEKRPLEINPAKFPAGASRFDIEGALGDCWLLAVASISGTFQFLDYVPPDQEL 180
 QY 132 Q-ENYAGIFHFQFQWGEWVVDRLPTKQD-ELLFVHSAEGSEFMSALLEKAVAKI 188
 Db 132 Q-ENYAGIFHFQFQWGEWVVDRLPTKQD-ELLFVHSAEGSEFMSALLEKAVAKI 188

Db 181 KGEYVGVVRRFRFGHWVEVLIDRLPVROGINTLVFMSNDPTFMSALKEKAYAKL 240

QY 189 NGCEYALSGATTEGDFETGGLAEWYEL--KKPPNLFKIIQKALQKSLGCSIDITS 246

Db 241 NGCYAHLSGGSQSASMDLITGGICLSLELNKERSPLIDOLKIYARCCIMGCSID--- 297

QY 247 AADSEAITFOK---LYKHAIVSYTGAEVENSGLCKLIRINPMGEV-EMTGRMNDNC 301

Db 298 ----SSVMEQKMDNGLIGSHAYSLTGYPVYVRRRTQMLMRPMPWDSHEMKAMDCGS 353

QY 302 PSMTNIDPEERERLRTNRH-EDGEFMSFSDLRHYSLEICNLTPDTL--TSQYKKNKL 358

Db 354 POREISEQEKKINLSPTADGEFMSYEDFCYFSRVVCHGLGLESLEYNQNHGKRL 413

QY 359 TK--MDGNMRGSTAGCCRRYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQHR 416

Db 414 DEAFSGQWRNVAVAGCINRRTYTNPQRIITVEDPDDDK--CSVLIGLMQTDI 471

QY 417 RRQRKMGEDMTTIGFGLYEVPEELSGQTNHLSKNPFLNRRERSPTFINLRVLRPK 476

Db 472 R--KKVQADPOPIGFVYVNAADDL---NTLSRAQLTSTRPIAKSQ--FINTREVTQAFR 524

QY 477 LPGEYILVSTPEPNNDGFCIRVFESEKADYAV-----DDEIEA-NLEEFDI 525

Db 525 VPEGSYVIPSITPDENIEVNFILKVFQTSITTEQLEDEDNTQGLPDVILEALKEDTTL 584

QY 526 SEDDIDDVRRFLAQLAGEDAE--ISAPEIQTLIRVLAARODIKS-DGFSLETCKIMV 581

Db 585 DE---DEIEQKF--LAIRPKTNAINAVKLGELINN--SLQDIPRFGPKNELCKSMV 637

QY 582 DMLDSGSGKLGKEFYITMTKIQKQKTYREIDVDSGTMSYEMKALBEAGFKMPCQ 641

Db 638 ASVDNNLTGVHEINFEFMDLWIOAKGMGHIFIKHVDQSGYFAVEFREALNDAGYHSNR 697

QY 642 LHQIVARFAD---DQILIDPNVRCVLETLTKFKQKDPEN-IGTIEBLDISLWLC 697

Db 698 LINAIINRYODPGTDK--ISFEDMLCMVRKLTAFETI-EAHPKNIEGTSLFSAEDLRF 754

QY 698 SV 699

Db 755 RV 756

RESULT 14

S44749

C06G4.2 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: S44749

R:Waterston, R.

A:Submitted to the EMBL Data Library, November 1993

A:Description: Sequence of the C. elegans cosmid C06G4.

A:Reference number: S44747

A/Accession: S44749

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-653 <WAT>

A:Cross-references: EMBL:L25598; NID:G409290; PIDN:AAA27940.1; PID:G409293

C:Genetics:

A:Insertions: 26/1; 69/1; 216/3; 269/3; 357/2; 401/2; 558/1; 583/3; 613/3

C:Superfamily: C06G4.2 protein; calpain catalytic domain homology

F/245-498/Domain: calpain catalytic domain homology <CALP>

Query Match 31.9%; Score 1184; DB 1; Length 653;

Best Local Similarity 47.3%; Pred. No. 8.3e-71;

Matches 231; Conservative 73; Mismatches 124; Indels 60; Gaps 9;

QY 30 QDYEALENECTLEAGTLFQDPSFPAIPALGFKELGYPSSKTRGMRKRPTIEICADPOFI 89

Db 205 QDPELEADQGLSEKRLPEFQFLANDSSLFSSKRP-----KRVEMLRPGEITREPLIT 259

QY 90 GGATRIDICGALGDCMLAALASLTNEETLARAVPLNOSPOENVAGIFHPQFOWYGEW 149

Db 260 EGHSRPDVITQGLGDCMLAAANLTCLKDELFYRVVPPDQSFTEYVAGIFHPQFOWYGEW 319

QY 150 VEYVVDRLPTKXGELLFVHASGESEFWSALKEAKYAKINGCYEALSGATTEGDFETG 209

Db 320 VDVIIDRLPTSGELLVTHASANNFWSALLKRAAKLFGSYEALKGTTSEALBDMTG 379

QY 210 GIAEWELKKPPNLFKIIQKALQKSLGCSIDITSAAADS---EAITFOKLYGHAYSV 266

Db 380 GLTEFIDLKPPRNLMQMMWRGFMGSLFQCSIE---ADPNWEAMSGVLKGHAYSI 435

QY 267 TGAEVENSNGSLCKLIRINPMG-EVMTGRMNDNCPSMTI--DPEERELTRHEDGER 324

Db 436 TGRIVDPNGQTCILIRINPMGNEBQWMPWSDNSREKMSVDSYKQDMGLFDDHGEF 495

QY 325 WMSFDFLRHYSRLICNLTPDITLSDTYKK-----WKTQKDNRRRSTAGGR 375

Db 496 WMSFDFMRHFEKMEICNLGPDYM-DEVYQMTGVKAAQVMAANTHDGANVRQTAGGR 554

QY 376 NYPTFMNPOYLKLEEBEDEDESGCTFLVGLIQHRRRQRKMGEDMTIGFGLYE 435

Db 555 NYITFANNPQFRVQLTSDPDDDD--ELCT----- 583

QY 436 VPEELSGQTNHLSKNPFLNRRERSDPTFINLRVLRPKLPGEYILVSTPEPNKDG 495

Db 584 ----AGNNRGRLSKQFPAKNSAMRSAAFINLREMTGRVRPBGVYVVPSTPEPNEA 638

QY 496 DFCIRVFS 503

Db 639 EFMRLRYT 646

RESULT 15

A24815

calpain (BC 3.4.22.17) large chain 1 - rabbit (fragments)

N:Alternate names: calcium-activated neutral proteinase (CANP)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 24-Jan-1988 #sequence_revision 24-Feb-1994 #text_change 16-Jul-1999

C/Accession: A24815; A41418

R:Emori, Y.; Kawasaki, H.; Sugihara, H.; Imajob, S.; Kawashima, S.; Suzuki, K.

J. Biol. Chem. 261, 9465-9471, 1986

A>Title: Isolation and sequence analyses of cDNA clones for the large subunits of two 1

A:Reference number: A92594; WUID:86250902; PMID:2424911

A:Molecule type: mRNA

A/Accession: A24815

A:Residues: 19-320 <EMO>

A:Cross-references: GB:M13363; NID:G165667; PIDN:AAA1456.1; PID:G165668

R:Kawasaki, H.; Imajob, S.; Suzuki, K.

J. Biochem. 102, 393-400, 1987

A>Title: Separation of peptides on the basis of the difference in positive charge: simul

A:Reference number: A41418; WUID:88032960; PMID:3667575

A/Accession: A41418

A:Molecule type: protein

A:Residues: 1-18;125-154;313-320 <KAW>

A>Note: Sequence was deduced from composition by homology

C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain; C/key words: blocked amino end; calcium binding; cysteine proteinase; duplication; EF hai

F/148-179/Domain: calmodulin repeat homology <EF1>

F/191-223/Domain: calmodulin repeat homology <EF2>

F/224-253/Domain: calmodulin repeat homology <EF3>

F/256-288/Domain: calmodulin repeat homology <EF4>

F/289-320/Domain: calmodulin repeat homology <EF5>

Query Match 24.9%; Score 925.5; DB 2; Length 320;

Best Local Similarity 56.6%; Pred. No. 4.5e-54;

Matches 167; Conservative 64; Mismatches 63; Indels 1; Gaps 1;

QY 402 ESGCTFLVGLIQHRRRQRKMGEDMTTIGFGLYEVPEELSGQTNHLSKNPFLTNRRAR 461

Db 20 ESGCSYVALMQGRRRRRRRFGDMETIGFAVYEVRELVGQALMLKDKDFLANSRAR 79

QY 462 SDTFINLRVLRNFKLPGEYILVSTPEPNKDGFCIRVFSKKADYQAVDDEIENL- 520

Db 80 SEQFINLRVSTFRPLPGEYVVPSTFEPNKGDEVLRFPSKRAQTQELDQIOANLP 139

APPLICANT: Hitachi, Ltd.
TITLE OF INVENTION: AIP1-INTERACTING PROTEINS AND USE THEREOF
FILE REFERENCE: HITA.0164
CURRENT APPLICATION NUMBER: US/10/370,481
PRIOR FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: US 60/358,369
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 60/019,808
PRIOR FILING DATE: 2002-09-03
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 21
LENGTH: 191
TYPE: PRT
ORGANISM: Homo sapiens
US-10-370-481-21

Query Match 4.5%; Score 168; DB 6; Length 191;
Best Local Similarity 25.0%; Pred. No. 2.7e-08;
Matches 41; Conservative 39; Mismatches 68; Indels 16; Gaps 3;

QY 553 LQTLRRVLA-KRQDIKSD-----GFSIETCKIWMVLDSDSGSKGLKEFYI 599
DB 28 LNVFQKVDKXRGVISTDELQALSNGTPTPFPVTKSIISMFDRNKAGVNFSEFTG 87
QY 600 LMTKIQKQKTYREIDVDKSGTMSYEMRKALBEGFMPQQLHGVYARF-ADDQLIID 658
DB 88 VMKVTDMQWVFRFYDRNSGMIDNKLQALSGGVYALSDQPHDILIRKDRGQGLA 147
QY 659 PDNVRCLVRLKTLTKIFKQDPENTGTIELDLSWL--CRSVL 700
DB 148 FDDPIQGGIVLQRLTDFRRYDTQDQWTVSYEYSLMVSIV 191

RESULT 3
US-60-479-073-498
Sequence 498; Application US/60479073
GENERAL INFORMATION:
APPLICANT: De Wilde, Gert Jules Hector
APPLICANT: Saunders, Michael John Scott
APPLICANT: Logghe, Marc Georges
TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
FILE REFERENCE: D00590.70042.US
CURRENT APPLICATION NUMBER: US/60/479,073
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 526
SOFTWARE: PatentIn version 3.2
SEQ ID NO 498
LENGTH: 5373
TYPE: PRT
ORGANISM: Homo sapiens
US-60-479-073-498

Query Match 3.2%; Score 117; DB 7; Length 5373;
Best Local Similarity 19.5%; Pred. No. 0.28;
Matches 114; Conservative 84; Mismatches 214; Indels 172; Gaps 24;

QY 180 LLEKAYAKINGCYALSGGATTEGFEFTGIAEWELKPPNLFKTIQALQKSLG 239
DB 2659 LEDLAADRINRLQAL--ASTQFOQMDELRTWLDK-----QSQAKNCPIIS 2705
QY 240 CSIDITSADSEATTFOKLVKGH--AVSYTGAE-----EVSNGSLQ-KLIRIRNP 287
DB 2706 AKLERLOSQLOENEBFQKSLNQHSGSYEVIVABGSLILSVPGEBKRTQONQVLELKH 2765
QY 288 WGE-----VEWTGRWNCPSMN-----TIDPERERLTLR 317
DB 2766 WEELSKTADRSQRLDCMQKQKQYQWVHEDLVPIWIDCKAKMELARTLDPVQLSSLL 2825
QY 318 RHEDGEFWMSPDFLRHYSRLICNLTPDTLTSDTYKKMKLTKMDGNWRGSLAGGCRNY 377

DB 2826 RSK-----AMLNEVERKRSLLILNSAAD----- 2849
QY 378 PNTFWMPQVLTLEBEDDEDESGCFVLGLIQHRRRQRMGEMHTIGFYIEVP 437
DB 2850 -----ILINSSADED-----G-----RDKAG-----INQNDAYT 2877
QY 438 EELSGQTNHLSKNFPLTNARERSDTFNLREVLNRFKLPGEYILVPSFEPNKGDF 497
DB 2878 EELQAKTG-SLEF---MTQLRBFQESFNKIEKVGAKQLRIFALGSAQCSNNK--- 2930
QY 498 CIRVFSKQADYQAVDEITANIEEF-DISEDDIDGVRFLPQALAGEDEIFAFLQTI 556
DB 2931 -----LEKRAQGEVLALEPPQVLYLNTFTQGLVEDAPDGSDSLHQAQAEVLE 2985
QY 557 LRRVLA--KRQDIKSDGFSIETCKI-----MYDMLDS-DGSGKLQ----- 593
DB 2986 KQVNSGVCMNEMKLEIGQFHCVRVEMPSQLDLDELDMGALIRGDTLSQSIQIEDVR 3045
QY 594 --LKEFYILMTKIQKTYREIDVDKSGTMSYEMRKALBEGFMPQQLHGVYARFA 651
DB 3046 LFLNKIHLKLDIEASBAECRHW-LBEGTLDLGLKRELEALN-KQCGKLTGRGKAR-- 3101
QY 652 DDQLIIDPDNVRCLVRLKTLTKIFKQDPENTGTIELDLSWL 695
DB 3102 QEQLEL-----TLGRVEDPYRKLGLNDATTAAEEAEALQWV 3138

RESULT 4
US-60-479-073-500
Sequence 500; Application US/60479073
GENERAL INFORMATION:
APPLICANT: De Wilde, Gert Jules Hector
APPLICANT: Saunders, Michael John Scott
APPLICANT: Logghe, Marc Georges
TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
FILE REFERENCE: D00590.70042.US
CURRENT APPLICATION NUMBER: US/60/479,073
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 526
SOFTWARE: PatentIn version 3.2
SEQ ID NO 500
LENGTH: 5430
TYPE: PRT
ORGANISM: Homo sapiens
US-60-479-073-500

Query Match 3.2%; Score 117; DB 7; Length 5430;
Best Local Similarity 19.5%; Pred. No. 0.28;
Matches 114; Conservative 84; Mismatches 214; Indels 172; Gaps 24;

QY 180 LLEKAYAKINGCYALSGGATTEGFEFTGIAEWELKPPNLFKTIQALQKSLG 239
DB 2715 LEDLAADRINRLQAL--ASTQFOQMDELRTWLDK-----QSQAKNCPIIS 2761
QY 240 CSIDITSADSEATTFOKLVKGH--AVSYTGAE-----EVSNGSLQ-KLIRIRNP 287
DB 2762 AKLERLOSQLOENEBFQKSLNQHSGSYEVIVABGSLILSVPGEBKRTQONQVLELKH 2821
QY 288 WGE-----VEWTGRWNCPSMN-----TIDPERERLTLR 317
DB 2822 WEELSKTADRSQRLDCMQKQKQYQWVHEDLVPIWIDCKAKMELARTLDPVQLSSLL 2881
QY 318 RHEDGEFWMSPDFLRHYSRLICNLTPDTLTSDTYKKMKLTKMDGNWRGSLAGGCRNY 377
DB 2882 RSK-----AMLNVEVERKRSLLILNSAAD----- 2905
QY 378 PNTFWMPQVLTLEBEDDEDESGCFVLGLIQHRRRQRMGEMHTIGFYIEVP 437
DB 2906 -----ILINSSADED-----G-----RDKAG-----INQNDAYT 2933
QY 438 EELSGQTNHLSKNFPLTNARERSDTFNLREVLNRFKLPGEYILVPSFEPNKGDF 497

```
Db      2934 BELOAKTG-SLEE---WTORLREFQESFKNIKKVEGAKQLEIFDALGQACSNKN--- 2986
Qy      498 CIRVSEKADYOAVDEIEANLEEF-DISEDDIDGVRRLFAQLAGEDAISAFAELQTI 556
Db      2987 -----LEKRAQOEVLQALPQVDYLRNFTQGLVEDAPDSQSLHQAQOEFLV 3041
Qy      557 LRRVLA-KRODKSDGFSIETCKI-----MVMDLDS-DQSGKLG----- 593
Db      3042 KORVNSGCVMMENKLEGIQGFHCVRFMPSQLADLDELQMGALIGRDTSLQSIEDVR 3101
Qy      594 --LKEFIILMTKIQKQKIRREIDVDRSGTMSYEMKALEEAGFKPCCQLHOVYARFA 651
Db      3102 LFNKHLVHLKDLLEASAEGRHM-LBEEGTLDLGLKRELALN-KCGKGLTERGAR-- 3157
Qy      652 DDOLIIDFNFVACLVRLETLFKIFKQJDEPENTGTIELDLISWL 695
Db      3158 QEOLEL-----TLGRVEDFYRKLGKLNATTAABEAALQNV 3194
```

```
RESULT 5
US-10-273-573-8670
; Sequence 8670, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ. ID NOS: 10994
; SOFTWARE: Custom
; SEQ. ID NO: 8670
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(259)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8670
```

```
Query Match      2.9%; Score 108.5; DB 6; Length 259;
Best Local Similarity 28.4%; Pred. No. 0.013;
Matches 42; Conservative 24; Mismatches 61; Indels 21; Gaps 7;
Qy      538 FAQAGEDAISAFAELQTIILRVLAARODIKSDG---FSIETCKIWMVMDLSDSGKLG 594
Db      61 FAVAVAGDQDIDDELQRCLE-----TQSGIAGGYKPFNLETCLRWVSMIDRMSTGMEF 114
Qy      595 K-EFYL-----WTKIQKQKIRREIDVDRSGTMSYEMKALEEAGF-KMPCCQLHOV 645
Db      115 SIEFKLLGLLEWAGROHHSFSTDRWEXQDPELPE--GPXQTMGEFVXVPCGXIS 172
Qy      646 IVARFADDOLIIDFNFVACLVRLETLF 673
Db      173 FAKRYSTNGK-ITFDDYIACCVQTXGVF 199
```

```
RESULT 6
US-10-280-962A-45
; Sequence 45, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT APPLICATION NUMBER: US/10/280,962A
; CURRENT FILING DATE: 2002-10-25
; NUMBER OF SEQ. ID NOS: 56
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 45
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Thermococcus sp.
US-10-280-962A-45
```

```
Query Match      2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;
```

```
Qy      22 ERAIKLNDYELARNECELAGTLFODPSFPAIPSLAKRELGPF-----SSKTRGR 74
Db      132 EEEKLKMSDIEETLYHGESEFGF-----GPIMISYADESEARVIT 172
Qy      75 WKR-----PREICADPOFIIGATRT-----DICQALGDCWMLAI-----ASLTL 116
Db      173 WKKIDLPYEVVSTEEKMTKRLRVKEDQPDVLTITNGDNPFALKRCCEGLSVSFTL 232
Qy      117 NEE-ILARVPLNQSFOENYAGIFHPQFMOYGEVVEVVDRLPTKDGELLFVHSAAGSE 175
Db      233 GRDSEPKIQMGDRFAVEYKGVHFDLYPV---IRRTIN--LFTYTLNAVY-----EA 281
Qy      176 FMSALLEKAYAKINGCYEALSGATTGEPEDFTGIAEW-----YELKPPNLFKI 227
Db      282 VFQKPKKXYAA-----EIAIAMEGEGLE---RVARYSMEDARVYETIGR-----EF 326
Qy      228 IQKALQKSLGCSI-DITSADSEAITFOQLVKHAYSVTGAEEVESNGSLQKLRIRN 286
Db      327 FPMEAQSLRILQGLMDVRS-----STGNLVEWFLRK 360
Qy      287 PMGEVEMTGRMNDNCPNNTIDPEERERLTRR-----HEDGEFW--MSFSDPLRHY 335
Db      361 AYER-----NELAN---KDERELARRGGVAGVYKEPERELMNIVYLDERSLY 409
Qy      336 SRLKIC-NLTPDTLSDTYTKMKLTMDGNMRGRTAGCCRNYPNTFMNPOYLKLEEB 394
Db      410 PSIIITHNVSPTLNE-----GCRSY----- 431
Qy      395 DEDEDESGSC-----TLVGLQKRRRQKMGEMHTTIGFIVEPEELSGQTN 446
Db      432 DVAPVEGHRKCPDPGFIPSLIGNLLEERQKIKRKRAKATLDPLEKNLDY-----RORAI 486
Qy      447 HLSKNFF-----LTNRARERSPTFINLREVLNFKLP----- 478
Db      487 KILANSYGYGYARARWYCRCAESVTAMGEYIEMV---RELEKRFKXLYADTDGL 544
Qy      479 ----PG-----EYI--LVSTFEPNKGDFCIRVSEKADYOAVDEIEA 518
Db      545 HATIPGADAETVKKAMEFLANINPCLPGLLELEYEG-FYVRGFFVTKKKYAVIDEBGKI 603
Qy      519 NLEFPDISDDIDDGVRRLFAQLAGEDAISAFAELQTIILRVLAARODIKSDGFSIETCK 578
Db      604 TTRGLEI-----VRDWSIA--KETQARVLEAILRH-----GDVEAVR 641
Qy      579 IWMVMDSDSGKLGKLEFYILMTKIQKQKIRREIDVDRSGTMSYEMKALEEAGFKM 638
Db      642 IYREVTEK-----LSKYEVEPEKLVHQBQITREIK-DYKATGEPVVALAKLAARGVAKI 693
Qy      639 -PCQLHOVIV---ARFADDOLIIDFNF-----VRCLVLETLFKIF--K 677
Db      694 RPTVVISYIVLKGSGRIGD--RAIPFDEDPPTKHKYADADYIENQVLPAVERILRAFGVR 751
Qy      678 QLDPENTGTIELDLISWL 695
Db      752 KEDLRYQKTRQVGLQAWL 769
```

```
RESULT 7
US-10-280-962A-46
; Sequence 46, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
```


FILE REFERENCE: 25436/2342
CURRENT APPLICATION NUMBER: US/10/280,962A
CURRENT FILING DATE: 2002-10-25
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46
LENGTH: 776
TYPE: PR
ORGANISM: Thermococcus sp.
US-10-280-962A-46

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

22 ERAIKYNDYELARNECLEAGTIFODPSFPAISALGFKELGY-----SKTRGMR 74
132 EEBLKMSFDIELTYHGEFRT-----GPIIMISYADESEARVIT 172
75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLTL 116
173 WKIDLPYEVVSTEEKMIRKPLRVKEDPDVLITNGDNFPAIYKRCCKGVSFTL 232
117 NEB-ILARVPLNOSFOENTAGIFHFQFOWGEWEVVDDRLPTKDGELLFVHSAEGSE 175
233 GRDSEPKIQMGDRFAVEKGRVHFDLYPV---IRRTIN--LPTVTLLEAVY-----EA 281
176 FMSALLEKAYAKINGCEALSGATTEGPFDTGSIAMW-----YELKPPPLFKI 227
282 VFQKPEKRYAE-----EIAWETGEGLR---RVARYSMEDARVTEYELGR-----EP 326
228 IQALQKSLGSGSI-DITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKLRIRN 286
327 FPMBAQSLRIQGLMDVRS-----STGNLVEMWFLRK 360
287 PMGEVWTCRWNDPCSWNTIDPEERELTR-----HEDGEFW--MSFSDPLRHY 335
361 AYER-----NELAPN---KPDRELARRGGVAGYVKEPERGLMDNIVYLDERSLY 409
336 SRLKIC-NLTPDTLSDTYKKMKLTMDGNMRRGSTAGGCRNYPNTFMNPOYLILBEE 394
410 PSIIITHNVSPTLNR-----GCRSY-----431
395 DEDEDESGC-----TFVLGLQKRRRQKMGEDMHTIGFGIYEVEBELSGQTN 446
432 DVAPVGHKFCQDPGFIPLSLGNLBEROKIKRKKATIDPLEKNLDY-----RORAI 486
447 HSKNFF-----LTNRARERSDTFINLREVLNRFKLP-----478
487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVLYADTDGL 544
479 -----PG-----EYI--LVSTFEPNKGDPCIRVSEKKADYQVANDIEIA 518
545 HATIPGADATVKKKAMEFLANTINPKLPGLLEBYEG-FYVRGFTVTKKTYAVIDECKI 603
519 NLEEFISDDIDDGVRRLFAOLAGDAEISAFELQTLIRRVLAARODIKSDGSJETCK 578
604 TTRGLEI-----VARDMSEIA--KETQARVLEAILRH-----GDVEANVR 641
579 IAWDMLDSGSGKLGKEFYILMTKIOKYKTYREIDVDRSGTMSYEMRKALEEAGFM 638
642 IYREVTEK-----LSKYEVPEPKLVIHQITRELK-DYKATGPHVALAKRLAARGVKI 693
639 -PCQLHQVIVY---AFADDQLIIDPDRN-----VRCVLVLETLFKIF--K 677
694 RFGTVISYIVLKSGSGRIGD--RAIPDEPDPTKHKYDADYIENQVLPVAVERTILRARGYR 751
752 KEDLRVYKTRQVGLGAWL 769

RESULT 8

US-10-298-680A-45
Sequence 45; Application US/10298680A
GENERAL INFORMATION:
APPLICANT: Sorce, Joseph A
TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
FILE REFERENCE: 25436/2345
CURRENT APPLICATION NUMBER: US/10/298,680A
CURRENT FILING DATE: 2002-11-18
PRIORITY APPLICATION NUMBER: US 10/280,962
PRIORITY FILING DATE: 2002-10-25
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 45
LENGTH: 776
TYPE: PR
ORGANISM: Thermococcus sp.
US-10-298-680A-45

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

22 ERAIKYNDYELARNECLEAGTIFODPSFPAISALGFKELGY-----SKTRGMR 74
132 EEBLKMSFDIELTYHGEFRT-----GPIIMISYADESEARVIT 172
75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLTL 116
173 WKIDLPYEVVSTEEKMIRKPLRVKEDPDVLITNGDNFPAIYKRCCKGVSFTL 232
117 NEB-ILARVPLNOSFOENTAGIFHFQFOWGEWEVVDDRLPTKDGELLFVHSAEGSE 175
233 GRDSEPKIQMGDRFAVEKGRVHFDLYPV---IRRTIN--LPTVTLLEAVY-----EA 281
176 FMSALLEKAYAKINGCEALSGATTEGPFDTGSIAMW-----YELKPPPLFKI 227
282 VFQKPEKRYAE-----EIAWETGEGLR---RVARYSMEDARVTEYELGR-----EP 326
228 IQALQKSLGSGSI-DITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKLRIRN 286
327 FPMBAQSLRIQGLMDVRS-----STGNLVEMWFLRK 360
287 PMGEVWTCRWNDPCSWNTIDPEERELTR-----HEDGEFW--MSFSDPLRHY 335
361 AYER-----NELAPN---KPDRELARRGGVAGYVKEPERGLMDNIVYLDERSLY 409
336 SRLKIC-NLTPDTLSDTYKKMKLTMDGNMRRGSTAGGCRNYPNTFMNPOYLILBEE 394
410 PSIIITHNVSPTLNR-----GCRSY-----431
395 DEDEDESGC-----TFVLGLQKRRRQKMGEDMHTIGFGIYEVEBELSGQTN 446
432 DVAPVGHKFCQDPGFIPLSLGNLBEROKIKRKKATIDPLEKNLDY-----RORAI 486
447 HSKNFF-----LTNRARERSDTFINLREVLNRFKLP-----478
487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVLYADTDGL 544
479 -----PG-----EYI--LVSTFEPNKGDPCIRVSEKKADYQVANDIEIA 518
545 HATIPGADATVKKKAMEFLANTINPKLPGLLEBYEG-FYVRGFTVTKKTYAVIDECKI 603
519 NLEEFISDDIDDGVRRLFAOLAGDAEISAFELQTLIRRVLAARODIKSDGSJETCK 578
604 TTRGLEI-----VARDMSEIA--KETQARVLEAILRH-----GDVEANVR 641
579 IAWDMLDSGSGKLGKEFYILMTKIOKYKTYREIDVDRSGTMSYEMRKALEEAGFM 638
642 IYREVTEK-----LSKYEVPEPKLVIHQITRELK-DYKATGPHVALAKRLAARGVKI 693
639 -PCQLHQVIVY---AFADDQLIIDPDRN-----VRCVLVLETLFKIF--K 677
694 RFGTVISYIVLKSGSGRIGD--RAIPDEPDPTKHKYDADYIENQVLPVAVERTILRARGYR 751

QY 678 QLPDENTGTIELDLISWL 695
DB 752 KEDLRKYQKTRQVGLGAWL 769

RESULT 9
US-10-298-680A-46
; Sequence 46, Application US/10298680A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2345
; CURRENT APPLICATION NUMBER: US/10/298,680A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US 10/280,962
; PRIORITY FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 776
; TYPE: PR1
; ORGANISM: Thermococcus sp.
US-10-298-680A-46

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

QY 22 ERAIKYLNODYEALNRECEAGTLFODPSFPALPSALGFKELGPT-----SSKTRGMR 74
DB 132 EEBELKUMSFDETLHYHGESEFGT-----GPILMISYADSEARVIT 172
QY 75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLT 116
DB 173 WKKIDLPYEVVSTEEKMTRFLRVYKVKOPDVLITNGDNFPAVKRCEKLVGSFTL 232
QY 117 NEB-ILARVPLNQSFOENYAGIFHPQMOYGEVVEVVDRLPTKDGELLFVHSAEGSE 175
DB 233 GRDSEPKIQMGDRFAVEYKGRVHFDLPV---IRRTIN--LPTTLEAVY-----EA 281
QY 176 FMSALLEKAYAKINGCEYALSGATTEGFEDFTGGIAEW-----YEIKRPPNLFKI 227
DB 282 VFGRPKPEKYYAE-----EIAIAMEYEGGLE---RVARYSMEDARVYELGR-----EF 326
QY 228 IQKALQKSLGSGSI-DITSADSEAITFOKLVKGHAIVTGAEEVENSGLQKLIRIN 286
DB 327 FPMBAQLSRLIGGLMDVSR-----STGNLVEMWFLLRK 360
QY 287 PMGEVEMTGRMNDNCPSMNTIDPEERERLTR-----HEDGEFW--MSFSDPLRHY 335
DB 361 AYER-----NELAPV---KDERELARRRGVAGYVKEPERGLMDNIIVLDPRSLY 409
QY 336 SRLKIC-NLTPDTLTSPTYKKMKLTMDGNMRGSGTAGCRNYPNTFMNPQYLIKLEEE 394
DB 410 PSIIITHNVSPTLNRE-----GCRSY----- 431
QY 395 DEDEDEGESGC-----TFVLGLIOKRRRQKMGEDMHTTIGFGLIYVPEELSGOTNI 446
DB 432 DVAPVGHKFCXDPGPFISLLGNLLEERQKIRKKKATLDPLEKULDY-----RQRAI 486
QY 447 HLSKNFF-----LTNRARSDPTFINLREVLNFKLP----- 478
DB 487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVUYADTDGL 544
QY 479 ----PG-----EYI--LVSTFEPNKDPCIRVFSKKADYQAVDDEIEA 518
DB 545 HATIPGADAEIVYKKAEMFLANINPKLPGLILEYEG--FYVGGFVTKKKYAVIDDEGKI 603
QY 519 NLBEFDISEDDIDGVARLFAOLAGDEAETISAFELQIILRLVLAQRDIDISDGSJETCK 578
DB 604 TTGLEI-----VARDWSEIA--KETQARVLEAILRH-----GDVEEAVR 641

QY 579 IMVMDLSDSGSKGLKEFYIIMTKIQKYQKTYREIDVRSCTMANSYEMRKALIEAGFKM 638
DB 642 IYREVEK-----LSKYEVPPEKLVIHQOLTRERK-DYKATGPVALAKRLAARGVKI 693
QY 639 -PCQAHQVY-----ARFADQOLIIDPDNF-----VRLGVNLETILFKI--K 677
DB 694 RFGTVISYIVLKSSGIGD--RAIPDEDPPTKHKYDADYIENQVLAVERILRAFGYR 751
QY 678 QLPDENTGTIELDLISWL 695
DB 752 KEDLRKYQKTRQVGLGAWL 769

RESULT 10
US-10-324-846-55
; Sequence 55, Application US/10324846
; GENERAL INFORMATION:
; APPLICANT: Bohns, Michael
; TITLE OF INVENTION: DNA Polymerase Blends and Uses Thereof
; FILE REFERENCE: 25436/2362
; CURRENT APPLICATION NUMBER: US/10/324,846
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 776
; TYPE: PR1
; ORGANISM: Pyrococcus GB-D
US-10-324-846-55

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

QY 22 ERAIKYLNODYEALNRECEAGTLFODPSFPALPSALGFKELGPT-----SSKTRGMR 74
DB 132 EEBELKUMSFDETLHYHGESEFGT-----GPILMISYADSEARVIT 172
QY 75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLT 116
DB 173 WKKIDLPYEVVSTEEKMTRFLRVYKVKOPDVLITNGDNFPAVKRCEKLVGSFTL 232
QY 117 NEB-ILARVPLNQSFOENYAGIFHPQMOYGEVVEVVDRLPTKDGELLFVHSAEGSE 175
DB 233 GRDSEPKIQMGDRFAVEYKGRVHFDLPV---IRRTIN--LPTTLEAVY-----EA 281
QY 176 FMSALLEKAYAKINGCEYALSGATTEGFEDFTGGIAEW-----YEIKRPPNLFKI 227
DB 282 VFGRPKPEKYYAE-----EIAIAMEYEGGLE---RVARYSMEDARVYELGR-----EF 326
QY 228 IQKALQKSLGSGSI-DITSADSEAITFOKLVKGHAIVTGAEEVENSGLQKLIRIN 286
DB 327 FPMBAQLSRLIGGLMDVSR-----STGNLVEMWFLLRK 360
QY 287 PMGEVEMTGRMNDNCPSMNTIDPEERERLTR-----HEDGEFW--MSFSDPLRHY 335
DB 361 AYER-----NELAPV---KDERELARRRGVAGYVKEPERGLMDNIIVLDPRSLY 409
QY 336 SRLKIC-NLTPDTLTSPTYKKMKLTMDGNMRGSGTAGCRNYPNTFMNPQYLIKLEEE 394
DB 410 PSIIITHNVSPTLNRE-----GCRSY----- 431
QY 395 DEDEDEGESGC-----TFVLGLIOKRRRQKMGEDMHTTIGFGLIYVPEELSGOTNI 446
DB 432 DVAPVGHKFCXDPGPFISLLGNLLEERQKIRKKKATLDPLEKULDY-----RQRAI 486
QY 447 HLSKNFF-----LTNRARSDPTFINLREVLNFKLP----- 478
DB 487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVUYADTDGL 544
QY 479 ----PG-----EYI--LVSTFEPNKDPCIRVFSKKADYQAVDDEIEA 518
DB 545 HATIPGADAEIVYKKAEMFLANINPKLPGLILEYEG--FYVGGFVTKKKYAVIDDEGKI 603

```

OY      NLEEFIDSEDDIDDGVRRFLPAQLAGDAEISAFEELOTTIRRYLAKRQDIXKDSGFSIETCK 578
Db      604 TTRGGEI-----VRDMSEIA--KETQARVLEALLRH-----GDVEEAYR 641
OY      579 IMVMDLSDSGSKGLGIKEFYIIMTKIQXYQKIYREIDVDRSGTMSYEMRKALEAGFYK 638
Db      642 IYREVTEK-----LSKYEVPPEKIVIEHQITRELK-DYKATGPHVALAKRLLAAGVGI 693
OY      639 -PCQHQYIV---ARFADQOLIIPDNH-----VRLVLETLTKIE-K 677
Db      694 RRGVYISIVLKGSGIGD-RAIPDEFDPTRKHYKDADYIENOVLPAYERILRAFGR 751
OY      678 QLDPENTGTEIDLISWL 695
Db      752 KEDLRKYKTRQVGLGAWL 769

RESULT 11
US-10-324-846-56
: Sequence 56, Application US/10324846
: GENERAL INFORMATION:
:   APPLICANT: Borns, Michael
:   TITLE OF INVENTION: DNA Polymerase Blends and Uses Thereof
:   FILE REFERENCE: 25436/2362
:   CURRENT APPLICATION NUMBER: US/10/324,846
:   CURRENT FILING DATE: 2002-12-20
:   NUMBER OF SEQ ID NOS: 113
:   SOFTWARE: PatentIn version 3.1
:   SEQ ID NO 56
:   LENGTH: 776
:   TYPE: PRT
:   ORGANISM: Pyrococcus GB-D
:   US-10-324-846-56

```

Query Match	2.8%	Score 105	DB 6	Length 776
Best Local Similarity	18.9%	Pred. No. 0.16		
Matches 151	Conservative 105	Mismatches 258	Indels 284	Gaps 40

```

0Y 22 ERAIYUONDYALNRNCEAGLTFQDPSPFPAISALGFKELGPY-----SSKTRGR 74
Db 132 EBEILKMSDITLTHGESEFGT-----GPIIMISVADESEARVIT 172
QY 75 WKR-----PEICADPOFIIGATRT-----DICQALDPCWLLAAI-----ASLTL 116
Db 173 WKKIDLPYVEVSTEXEMIKRFLRVYKEXPDLVLTYNNGDNFDPAYLKKRCEKLGVSFTL 232
QY 117 NEE-ILARVPLNQSFOENTYAGIFHHQFMOYGEWVEVVDRLPTKDELLEFVHSABGE 175
Db 233 GRDGEPEKIORMGDRFAVEKRVHVDLPV-----IRRTIN--LPTVTLLEVY-----EA 281
QY 176 FMSALLEXAYAKINGCYEALSGATGEGEDFTGGIAEW-----YEAKKPPNLFKI 287
Db 282 VFGKREKEKYAE-----ELATAMETEGLE---RPARSMEDARPTYELGR-----EF 326
QY 228 IOKALOKOSLIGCSI-DITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLOKLIRIN 286
Db 327 FPEMAQLSRLLIQGLMDVRS-----STGNLVEMFWFLRK 360
QY 287 PWGEVEMTGRKMNDCPSMNTIDPEERELTR-----HEDGEFW--MSSSDPLRHY 335
Db 361 AYER-----NELAPN---KPERELARRRGVAGVYKEPERGLMNDIVILDFRSLY 409
QY 336 SRLIEIC-NLTPRTISDYTKKMWLTMDGNWRRGSAAGCCRNYPNTFMNPOVLLKLEEE 394
Db 410 PSIIITTHNVSPTLNR-----GGRSY-----431
QY 395 DEDEDEGEGC-----TFLVGLIOKHRRROKMGEDMHTIGFGIYEVPELSGOTNI 446
Db 432 DVAREVGHNFCDPFPFISLJGNLLEBRQKIRKKKATILDPLEKNULDY-----RQRAI 486
QY 447 HUSKQPF-----LTNPARERSDPTINLREVLNRFKLP-----478

```

```

Db      487  KILANSYGYGYABARAPWCECAESVTAMGREYIEMWY--RELEEKGFPRVYADPDGL 544
QY      473  ---PG-----EYI--LVSTSEPKKODDFCIARVSEKKADYQAVDEIEA 518
Db      545  HATIPDAAEYKKAAMEFLNINPKLGGILLEBYEG--FYRGFFVTKKAYAVDEGKI 603
QY      519  NLEEPSIEDDIDGVRRLFAQLAGEDAEISAFELQTLIRRLAKRODIXSDGSIETCK 578
Db      604  TTRGAEI-----VRRDMSRHA--KETQARVLEALIRH-----GDVEEAVR 641
QY      579  IAMDMLDSDSGSKLGLKEFYILMTYQCKYVEIDVRSGTMSYEMERKALEEGFKM 638
Db      642  IAEVETEK-----LSKIEVPEPEKLVTHQITRELK--DYKATQPHYAIAKRLAARGVKI 693
QY      639  -PCQLHQVIV---ARFADQLIIDPNF-----VRCVRLSETLEFKIF--K 677
Db      694  RPTGVISYIVLKGSGRIGD--RAIPDFEDFDPTKHKADYADYIENQVPAVERILIRAGYR 751
QY      678  QLDPENTGTEIIDLISWL 695
Db      752  KEDLRVQKTRVGSLGAWL 769

```

```

RESULT 12
US-10-280-962A-41
; Sequence 41, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorce, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT APPLICATION NUMBER: US/10/280,962A
; CURRENT FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Pyrococcus sp.
US-10-280-962A-41

```

Query Match	2.8% ; Score 104.5 ; DB 6 ; Length 774 ;
Best Local Similarity	19.7% ; Pred. No. 0.17 ;
Matches 131 ;	Conservative 104 ; Mismatches 238 ; Indels 191 ; Gaps 33 ;
Qy	149 WVEVVDNR-----LPTKGELLFVHSAGSEFWASALLEKAVAKINGCYBALSGGA 199
Db	180 YVDVSTEREKIRKFLRVKKEKEDVDVITYGDNFDF--ATLKRCCKL-GINFALGRDG 236
Qy	200 TTSEFEDFTGIGIAFWELK-KRPNNLKTIQKALQKSLGCSLIDITSAADESAITPCKL 258
Db	237 SEPKIQKMGDRFA--VEVKGRIHDFLVIRKTNLP-----YTLAEYEAVPQPK 287
Qy	259 VKGHAVSVTGAEEVESN-----GSLQKLRIRNPWGEVE 292
Db	288 EKVAEETITPAWETIGENILERVARYSMEDAKVTYELGKFLPMEAQSLRI-----GQSL 341
Qy	293 W-----TG--RW-----NDNCPSWMTIDPERERLTRHEDGE-----F 324
Db	342 WDVARSTGMLVEFLLRKARERELAPN--KPEDEK-LARRQSYEGGYKPEBGL 396
Qy	325 W--WSPFDPLRHYSRLIC-NLPEDTLTSDTYKMKMLTKMDGNRRGSTAGACNNPYTF 381
Db	397 WENIVYIDFRLSYIIITINNVSDTLNRECKRYDVAPOUGH-----RFCDDPQFI 449
Qy	382 WNNFOYL-IKLEEDDEDGEGSCGPLVGLIQHRRRQRMGMDMTIGG----- 432
Db	450 ---PSLLGDLLEEROKIKKKAKATIDIERGLDYORAIKITLANSYYGYGVARAWYC 506
Qy	433 -----IYVEPEPLSGQTNHLSKNPFLTNRARERDTEFINREVLRN 474
Db	507 KECAESTYAMGREYITMTIKEIEKYGKAVYISDTGDFALITPGADAEIVKKKAMEFIN- 565
Qy	475 PKLPGEYI--LVSTEEPKNKGDFCIRVFSEKKADYOAVDIEANLEPDISDDIDD 532

Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKYAVIDESEKITTRGLEI----- 610
QY 533 GVRRLFAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGSKL 592
Db 611 -VRDMSEIA---KETQARVLBALK-----DGVKAVRIVKEVTEK----- 649
QY 593 GLKEFYILMTKIOKQIYREIDVDSGTMSYEMKALBEAGFKM-PCQLHGVIV---- 647
Db 650 -LSKYEVPPEKLVHQBQITRDLK-DYKATGPHVAVKRLAARGVKIRPGTVISYIVLKS 707
QY 648 ARPADQLIIDPFNF-----VRCLVLETLFKIF--KQDPENGTIYLDL 691
Db 708 GRIGD--RAIPDFEPTKHKYDAEYIENQVLAVERILRAFGYRKEDIRYQKTRQVGL 765
QY 692 ISWL 695
Db 766 SAWL 769

RESULT 13

US-10-280-962A-42
; Sequence 42, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorage, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US/10/280, 962A
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 774
; TYPE: PRN
; ORGANISM: Pyrococcus sp.
US-10-280-962A-42

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

QY 149 WVEYVVDNR-----LPTKDGELLFVHSAEGSEFWALLEKAYAKINGCYEALSOGA 199
Db 180 YDVVSTEREMIKRFLRVKEXKDPDLITYNGDNFDF--AYLKRCCEKL-GINFALGRDG 236
QY 200 TTGGEFDFTGGLAEWELEK-KPPNLEKTIIOKALQKSGSLGCSIDITSADSEAITFOKL 258
Db 237 SEPKIQMGDRFA--VEVKGRIHFDLYPVIRRTINLPT-----YTLBAVYEAVFGQPK 287
QY 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNPMGEVE 292
Db 288 EKYYABEITPAMWTGENLERVARYSMEDAKVYELGKEFLPMEQSLRLI-----GQSL 341
QY 293 W-----TG--RW-----NDNCPSWNTIDPEERELTRRHEDGE-----F 324
Db 342 WDVSRSTGNLVEMFLRKAYERNELAPN---KPEKE-LARRQSYEGGYKEBERGL 396
QY 325 W--MSFSDFLRHSRLKIC-NLTPDLTSDTYKKMKLTKMDGMMRGSTAGGRNPNYF 381
Db 397 WENIVYLDPRSLPSTIITHNVSPTLNREGCKEYDVAPOVGH-----RFCKDPGFT 449
QY 382 WNNPOYL-ITLBEDEDEBDESGCTFLVGLIQKHRRORRKMGEDEMTTIGFG----- 432
Db 450 ---PSILGDLLEBRQKIKKKMKATIDPIERKLDYQRAIKILANSYGYGYARARWYC 506
QY 433 -----IYVPEELSGQTNILHSKNFPLTNRARSSTPFINLREVLNR 474
Db 507 KECAESVTAMGREYITMTIKEIEEKYGFVYISDTPGFATIPGADAEYVKKKAMEFLN- 565
QY 475 FKLPPEYI--LVPSFTEPNKDGDFCIRVSEKKADYQAVDEIEANLEBFDISEDDIDD 532
Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKYAVIDESEKITTRGLEI----- 610

QY 533 GVRRLFAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGSKL 592
Db 611 -VRDMSEIA---KETQARVLBALK-----DGVKAVRIVKEVTEK----- 649
QY 593 GLKEFYILMTKIOKQIYREIDVDSGTMSYEMKALBEAGFKM-PCQLHGVIV---- 647
Db 650 -LSKYEVPPEKLVHQBQITRDLK-DYKATGPHVAVKRLAARGVKIRPGTVISYIVLKS 707
QY 648 ARPADQLIIDPFNF-----VRCLVLETLFKIF--KQDPENGTIYLDL 691
Db 708 GRIGD--RAIPDFEPTKHKYDAEYIENQVLAVERILRAFGYRKEDIRYQKTRQVGL 765
QY 692 ISWL 695
Db 766 SAWL 769

RESULT 14

US-10-298-680A-41
; Sequence 41, Application US/10298680A
; GENERAL INFORMATION:
; APPLICANT: Sorage, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2345
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US 10/280, 962
; PRIOR FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 774
; TYPE: PRN
; ORGANISM: Pyrococcus sp.
US-10-298-680A-41

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

QY 149 WVEYVVDNR-----LPTKDGELLFVHSAEGSEFWALLEKAYAKINGCYEALSOGA 199
Db 180 YDVVSTEREMIKRFLRVKEXKDPDLITYNGDNFDF--AYLKRCCEKL-GINFALGRDG 236
QY 200 TTGGEFDFTGGLAEWELEK-KPPNLEKTIIOKALQKSGSLGCSIDITSADSEAITFOKL 258
Db 237 SEPKIQMGDRFA--VEVKGRIHFDLYPVIRRTINLPT-----YTLBAVYEAVFGQPK 287
QY 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNPMGEVE 292
Db 288 EKYYABEITPAMWTGENLERVARYSMEDAKVYELGKEFLPMEQSLRLI-----GQSL 341
QY 293 W-----TG--RW-----NDNCPSWNTIDPEERELTRRHEDGE-----F 324
Db 342 WDVSRSTGNLVEMFLRKAYERNELAPN---KPEKE-LARRQSYEGGYKEBERGL 396
QY 325 W--MSFSDFLRHSRLKIC-NLTPDLTSDTYKKMKLTKMDGMMRGSTAGGRNPNYF 381
Db 397 WENIVYLDPRSLPSTIITHNVSPTLNREGCKEYDVAPOVGH-----RFCKDPGFT 449
QY 382 WNNPOYL-ITLBEDEDEBDESGCTFLVGLIQKHRRORRKMGEDEMTTIGFG----- 432
Db 450 ---PSILGDLLEBRQKIKKKMKATIDPIERKLDYQRAIKILANSYGYGYARARWYC 506
QY 433 -----IYVPEELSGQTNILHSKNFPLTNRARSSTPFINLREVLNR 474
Db 507 KECAESVTAMGREYITMTIKEIEEKYGFVYISDTPGFATIPGADAEYVKKKAMEFLN- 565
QY 475 FKLPPEYI--LVPSFTEPNKDGDFCIRVSEKKADYQAVDEIEANLEBFDISEDDIDD 532
Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKYAVIDESEKITTRGLEI----- 610
QY 533 GVRRLFAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGSKL 592

Db 611 -VRDWSA---KETOQAVLEALK-----DGDVEKAVRIKVEYTER----- 649
Qy 593 GLKEFYILMTKIQKYOKIYREIDVDRSGTNNSEMRKALEBAGFKM-PCQLHQVIV----- 647
Db 650 -LSKYEVPPEKLVHESQITRDLK-DYKATGPHVAVAKRLAARGVKIRPGIVISYIVLKS 707
Qy 648 ARPADQOLIIDFNF-----VRCLVLETLEFKIF--KQDPENTGTIELDL 691
Db 708 GRID--RAIPFDEPDTKHKYDAEYIENQVLPAYERILRAFGRKEDLRYQKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

RESULT 15
US-10-298-680A-42
Sequence 42, Application US/10298680A
GENERAL INFORMATION:
APPLICANT: Sorige, Joseph A
TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
FILE REFERENCE: 23436/2345
CURRENT APPLICATION NUMBER: US/10/298, 680A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US 10/280,962
PRIOR FILING DATE: 2002-10-25
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 42
LENGTH: 774
TYPE: PRT
ORGANISM: Pyrococcus sp.
US-10-298-680A-42

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

Qy 149 MVEVVVNDR-----LPTKDGELFVHSAEGSEFWSALLEKAYAKINGCYEALSGGA 199
Db 180 YVDVSTEREIMKRELVVKEKEDPDVLITNGDNDF--AYLKRCERL-GINFALGRDG 236
Qy 200 TTEGFEDFTGGIAEWEYLK-KPPNLFKTIQALQKSLGCSIDITSADESEAITFOKL 258
Db 237 SEPKIQRMGDRFA--VEYKGRHFDLPYVIRITNLP-----YTLEAYEAVFGQPK 287
Qy 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNFWGEVE 292
Db 288 EKVAEETTPAMETGENLERVARYSMEDAKTYELGKEFLPMEAOISRLI-----GQSL 341
Qy 293 W-----TG---RW-----NDNCPSWNTIDPEERERLRRHEDG-----F 324
Db 342 WDVSRSSTGNLVEWMLRKAYERNELAPN---KPEKE-LARRQSYEGGYKXEPERGL 396
Qy 325 W--MSFSDFLHYSLEIC-NLTPDTLSDTYKKWKLTKMDGNMRSGTAGCENYPTF 381
Db 397 WENIVYDPRSLYPSIITTHVSPDLNREGCKEYDVAPOYGH-----RCKDFPGFI 449
Qy 382 WNPQYL-IKLEEDDEDEDESCTFLVGLIQKHRRRQRMKGEDMTTIGF----- 432
Db 450 ---PSLGDLLERQKIKKKMKATIDPIERKLDYRQRAIKILANSYGYGYAARWYC 506
Qy 433 -----IYVPELSGOTNIHLSKNPFLTNRAERSDTPINLREVLR 474
Db 507 KECASVTAMGREYITMTIKEIIEKYGRVIYSDTGFPATLPGADAEIVKKKAMEFLN- 565
Qy 475 FKLPGEXY--LVPEFPNKDGRFCIRVSEKKADYQAVDDEIRANLEEFDISDID 532
Db 566 -----YINAKLPALALEYEG--FYKGFYPTKKKAVIIBEGKITRGLEI----- 610
Qy 533 GVRRLFAQLAGEDAEISAFELQTLIRVLARQDIKSDGFSIETCKIMVMDLSDSGKL 592

Db 611 -VRDWSA---KETOQAVLEALK-----DGDVEKAVRIKVEYTER----- 649
Qy 593 GLKEFYILMTKIQKYOKIYREIDVDRSGTNNSEMRKALEBAGFKM-PCQLHQVIV----- 647
Db 650 -LSKYEVPPEKLVHESQITRDLK-DYKATGPHVAVAKRLAARGVKIRPGIVISYIVLKS 707
Qy 648 ARPADQOLIIDFNF-----VRCLVLETLEFKIF--KQDPENTGTIELDL 691
Db 708 GRID--RAIPFDEPDTKHKYDAEYIENQVLPAYERILRAFGRKEDLRYQKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

Search completed: July 24, 2003, 13:13:09
Job time : 11.7077 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:51:55 ; Search time 422.806 Seconds

(without alignments)
1440.970 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MGAIAAKAKDRRAEGLS.....PENTGTIELDLISWLCFSVL 700

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 segs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/prodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/prodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/prodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/prodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/prodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/prodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/prodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/prodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/prodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/prodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/prodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/prodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/prodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/prodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/prodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/prodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/prodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/prodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/prodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/prodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/prodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/prodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/prodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/prodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/prodata/1/paa/US105_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3713	100.0	700	1	PCT-US02-10419-12
2	3713	100.0	700	16	US-09-281-253-12

3	3713	100.0	700	17	US-09-301-274-9	Sequence 9, Appl
4	3713 <th>100.0</th> <th>700</th> <th>19</th> <th>US-09-538-092-947</th> <th>Sequence 947, App</th>	100.0	700	19	US-09-538-092-947	Sequence 947, App
5	3713 <th>100.0</th> <th>700</th> <th>19</th> <th>US-09-562-979-9</th> <th>Sequence 9, Appl</th>	100.0	700	19	US-09-562-979-9	Sequence 9, Appl
6	3713 <th>100.0</th> <th>700</th> <th>22</th> <th>US-09-768-877-23</th> <th>Sequence 23, Appl</th>	100.0	700	22	US-09-768-877-23	Sequence 23, Appl
7	3713 <th>100.0</th> <th>700</th> <th>22</th> <th>US-09-791-537-8310</th> <th>Sequence 81010, A</th>	100.0	700	22	US-09-791-537-8310	Sequence 81010, A
8	3713 <th>100.0</th> <th>700</th> <th>23</th> <th>US-09-840-707A-9</th> <th>Sequence 9, Appl</th>	100.0	700	23	US-09-840-707A-9	Sequence 9, Appl
9	3713 <th>100.0</th> <th>700</th> <th>23</th> <th>US-09-840-707B-9</th> <th>Sequence 9, Appl</th>	100.0	700	23	US-09-840-707B-9	Sequence 9, Appl
10	3713 <th>100.0</th> <th>700</th> <th>23</th> <th>US-09-884-319A-7</th> <th>Sequence 7, Appl</th>	100.0	700	23	US-09-884-319A-7	Sequence 7, Appl
11	3713 <th>100.0</th> <th>700</th> <th>26</th> <th>US-10-038-557A-9</th> <th>Sequence 7, Appl</th>	100.0	700	26	US-10-038-557A-9	Sequence 7, Appl
12	3713 <th>100.0</th> <th>700</th> <th>27</th> <th>US-10-116-519-12</th> <th>Sequence 12, Appl</th>	100.0	700	27	US-10-116-519-12	Sequence 12, Appl
13	3713 <th>100.0</th> <th>700</th> <th>28</th> <th>US-10-219-051B-12485</th> <th>Sequence 12485, A</th>	100.0	700	28	US-10-219-051B-12485	Sequence 12485, A
14	3713 <th>100.0</th> <th>700</th> <th>28</th> <th>US-10-293-017-78</th> <th>Sequence 78, Appl</th>	100.0	700	28	US-10-293-017-78	Sequence 78, Appl
15	3713 <th>100.0</th> <th>700</th> <th>28</th> <th>US-10-293-071-78</th> <th>Sequence 78, Appl</th>	100.0	700	28	US-10-293-071-78	Sequence 78, Appl
16	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-296-180-112</th> <th>Sequence 12, Appl</th>	100.0	700	31	US-10-296-180-112	Sequence 12, Appl
17	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-296-180-112</th> <th>Sequence 12, Appl</th>	100.0	700	31	US-10-296-180-112	Sequence 12, Appl
18	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-300-620-12</th> <th>Sequence 4, Appl</th>	100.0	700	31	US-10-300-620-12	Sequence 4, Appl
19	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-350-064-4</th> <th>Sequence 390, App</th>	100.0	700	31	US-10-350-064-4	Sequence 390, App
20	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-389-987-390</th> <th>Sequence 390, App</th>	100.0	700	31	US-10-389-987-390	Sequence 390, App
21	3713 <th>100.0</th> <th>700</th> <th>31</th> <th>US-10-412-418-390</th> <th>Sequence 2, Appl</th>	100.0	700	31	US-10-412-418-390	Sequence 2, Appl
22	3705 <th>99.8</th> <th>700</th> <th>21</th> <th>US-09-736-860-2</th> <th>Sequence 7, Appl</th>	99.8	700	21	US-09-736-860-2	Sequence 7, Appl
23	3705 <th>99.8</th> <th>700</th> <th>21</th> <th>US-09-736-860-7</th> <th>Sequence 9, Appl</th>	99.8	700	21	US-09-736-860-7	Sequence 9, Appl
24	3705 <th>99.8</th> <th>700</th> <th>21</th> <th>US-09-736-860-9</th> <th>Sequence 63859, A</th>	99.8	700	21	US-09-736-860-9	Sequence 63859, A
25	3705 <th>99.8</th> <th>700</th> <th>22</th> <th>US-09-791-537-63859</th> <th>Sequence 63859, A</th>	99.8	700	22	US-09-791-537-63859	Sequence 63859, A
26	3698 <th>99.6</th> <th>699</th> <th>31</th> <th>US-09-791-537-65839</th> <th>Sequence 96, Appl</th>	99.6	699	31	US-09-791-537-65839	Sequence 96, Appl
27	3698 <th>99.6</th> <th>699</th> <th>31</th> <th>US-09-791-537-65839</th> <th>Sequence 2373, Ap</th>	99.6	699	31	US-09-791-537-65839	Sequence 2373, Ap
28	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-443-556-2373</th> <th>Sequence 23239, Ap</th>	98.7	701	31	US-09-443-556-2373	Sequence 23239, Ap
29	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-452-680-23239</th> <th>Sequence 14515, A</th>	98.7	701	31	US-09-452-680-23239	Sequence 14515, A
30	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-453-050-14515</th> <th>Sequence 14515, A</th>	98.7	701	31	US-09-453-050-14515	Sequence 14515, A
31	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-453-135-14515</th> <th>Sequence 14515, A</th>	98.7	701	31	US-09-453-135-14515	Sequence 14515, A
32	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-455-444-7816</th> <th>Sequence 7816, Ap</th>	98.7	701	31	US-09-455-444-7816	Sequence 7816, Ap
33	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-465-241-7816</th> <th>Sequence 14515, A</th>	98.7	701	31	US-09-465-241-7816	Sequence 14515, A
34	3666.5 <th>98.7</th> <th>701</th> <th>31</th> <th>US-09-466-412-14515</th> <th>Sequence 81227, A</th>	98.7	701	31	US-09-466-412-14515	Sequence 81227, A
35	3657 <th>98.5</th> <th>754</th> <th>21</th> <th>US-09-724-678-81227</th> <th>Sequence 81227, A</th>	98.5	754	21	US-09-724-678-81227	Sequence 81227, A
36	3657 <th>98.5</th> <th>754</th> <th>21</th> <th>US-09-724-678-81227</th> <th>Sequence 122174, A</th>	98.5	754	21	US-09-724-678-81227	Sequence 122174, A
37	3522 <th>94.9</th> <th>700</th> <th>22</th> <th>US-09-791-537-122174</th> <th>Sequence 4, Appl</th>	94.9	700	22	US-09-791-537-122174	Sequence 4, Appl
38	3518 <th>94.7</th> <th>700</th> <th>21</th> <th>US-09-736-860-4</th> <th>Sequence 8, Appl</th>	94.7	700	21	US-09-736-860-4	Sequence 8, Appl
39	3518 <th>94.7</th> <th>700</th> <th>21</th> <th>US-09-736-860-8</th> <th>Sequence 131857, A</th>	94.7	700	21	US-09-736-860-8	Sequence 131857, A
40	3518 <th>94.7</th> <th>700</th> <th>22</th> <th>US-09-791-537-91444</th> <th>Sequence 91444, A</th>	94.7	700	22	US-09-791-537-91444	Sequence 91444, A
41	3518 <th>94.7</th> <th>700</th> <th>22</th> <th>US-09-791-537-131857</th> <th>Sequence 131857, A</th>	94.7	700	22	US-09-791-537-131857	Sequence 131857, A
42	3518 <th>94.7</th> <th>700</th> <th>31</th> <th>US-09-350-064-12</th> <th>Sequence 12, Appl</th>	94.7	700	31	US-09-350-064-12	Sequence 12, Appl
43	3511 <th>94.6</th> <th>700</th> <th>22</th> <th>US-09-791-537-99714</th> <th>Sequence 99714, A</th>	94.6	700	22	US-09-791-537-99714	Sequence 99714, A
44	3417 <th>92.0</th> <th>691</th> <th>21</th> <th>US-09-724-678-81234</th> <th>Sequence 81234, A</th>	92.0	691	21	US-09-724-678-81234	Sequence 81234, A
45	3417 <th>92.0</th> <th>691</th> <th>21</th> <th>US-09-724-678A-81234</th> <th>Sequence 81234, A</th>	92.0	691	21	US-09-724-678A-81234	Sequence 81234, A

ALIGNMENTS

RESULT 1
PCT-US02-10419-12
Sequence 12, Application PC/TUS0210419
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
FILE REFERENCE: D0124 PCT
CURRENT APPLICATION NUMBER: PCT/US02/10419
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US 60/281,253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288,768
PRIOR FILING DATE: 2001-05-04
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/296,180
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/300,620
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-10419-12

Query Match 100.0%; Score 3713; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGIAAKIAKORAEAGISGHERAIKYINODYALNECELAGTLPQDSEFPALPSALGF 60
DB 1 MAGIAAKIAKORAEAGISGHERAIKYINODYALNECELAGTLPQDSEFPALPSALGF 60
QY 61 KELGPSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAALASLTINEEI 120
DB 61 KELGPSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAALASLTINEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSABSEFWASAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSABSEFWASAL 180
QY 181 LEKAYAKINGCYEALSGGATTEGFEPTGGIAEWELKRPPLFKIIQKALQKSLIGC 240
DB 181 LEKAYAKINGCYEALSGGATTEGFEPTGGIAEWELKRPPLFKIIQKALQKSLIGC 240
QY 241 SIDITSAADSEALTTFQKLVKHAASVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
DB 241 SIDITSAADSEALTTFQKLVKHAASVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
QY 301 CPSMNTIDPEERERLTRRHEDGFFMMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGFFMMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
QY 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBEGSGCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBEGSGCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRAERSDPTINLEVLNRFKLP 480
DB 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRAERSDPTINLEVLNRFKLP 480
QY 481 EYILVSTFEENKOGDFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYILVSTFEENKOGDFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKEFYIL 600
QY 601 WTKIQOKYKRIEIDVRSGMTNSYEMRKALBEAGFMPQOLHOIVARFADQLIIDPD 660
DB 601 WTKIQOKYKRIEIDVRSGMTNSYEMRKALBEAGFMPQOLHOIVARFADQLIIDPD 660
QY 661 NFVRCVRLTETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
```

RESULT 2

US-09-281-253-12
; Sequence 12, Application US/09281253
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124
; CURRENT APPLICATION NUMBER: US/09/281,253
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-281-253-12

Query Match 100.0%; Score 3713; DB 16; Length 700;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGIAAKIAKORAEAGISGHERAIKYINODYALNECELAGTLPQDSEFPALPSALGF 60
DB 1 MAGIAAKIAKORAEAGISGHERAIKYINODYALNECELAGTLPQDSEFPALPSALGF 60
QY 61 KELGPSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAALASLTINEEI 120
DB 61 KELGPSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAALASLTINEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSABSEFWASAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSABSEFWASAL 180
QY 181 LEKAYAKINGCYEALSGGATTEGFEPTGGIAEWELKRPPLFKIIQKALQKSLIGC 240
DB 181 LEKAYAKINGCYEALSGGATTEGFEPTGGIAEWELKRPPLFKIIQKALQKSLIGC 240
QY 241 SIDITSAADSEALTTFQKLVKHAASVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
DB 241 SIDITSAADSEALTTFQKLVKHAASVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
QY 301 CPSMNTIDPEERERLTRRHEDGFFMMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGFFMMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
QY 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBEGSGCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBEGSGCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRAERSDPTINLEVLNRFKLP 480
DB 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRAERSDPTINLEVLNRFKLP 480
QY 481 EYILVSTFEENKOGDFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYILVSTFEENKOGDFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKEFYIL 600
QY 601 WTKIQOKYKRIEIDVRSGMTNSYEMRKALBEAGFMPQOLHOIVARFADQLIIDPD 660
DB 601 WTKIQOKYKRIEIDVRSGMTNSYEMRKALBEAGFMPQOLHOIVARFADQLIIDPD 660
QY 661 NFVRCVRLTETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
```

RESULT 3

US-09-301-274-9
; Sequence 9, Application US/09301274
; GENERAL INFORMATION:
; APPLICANT: Fredking, Terry M.
; APPLICANT: Ignat'yev, George M.
; APPLICANT: ANTIBODY SYSTEMS, INC.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; FILE REFERENCE: 24881-0301
; CURRENT APPLICATION NUMBER: US/09/301,274
; CURRENT FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: comprising amino acid sequence

;; PUBLICATION INFORMATION:
;; PATENT DOCUMENT NUMBER: 5,817,476
;; PATENT FILING DATE: 1995-06-07
;; PUBLICATION DATE: 1998-10-06
US-09-301-274-9

Query Match 100.0%; Score 3713; DB 17; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAGIAATLADREABEGLSHERAIKYLNDYELARNCEAGTLFDDPSFPAIPSLGF 60
DB 1 MAGIAATLADREABEGLSHERAIKYLNDYELARNCEAGTLFDDPSFPAIPSLGF 60
QY 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIASTLTNEEI 120
DB 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIASTLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFWQYGEWVEVVDRLPTKGBELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFWQYGEWVEVVDRLPTKGBELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYELKKRPPNLFKIIQKALQKSLG 240
DB 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYELKKRPPNLFKIIQKALQKSLG 240
QY 241 SIITSAADSEALTFOKLVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIITSAADSEALTFOKLVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEBRERLTRRHEDGEFWSFSDFLRHYSRLICNLTPDTLTSPTYKKMLTK 360
DB 301 CPSWNTIDPEBRERLTRRHEDGEFWSFSDFLRHYSRLICNLTPDTLTSPTYKKMLTK 360
QY 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBDEDEBEGSCCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBDEDEBEGSCCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEBELSGQTNHLSKNFLTNARERSDPTFLNREVLNRPFKLP 480
DB 421 KMGEDMHTIGFGIYEVEBELSGQTNHLSKNFLTNARERSDPTFLNREVLNRPFKLP 480
QY 481 EYILVSTFEPNKGDFCIRVSEKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKGDFCIRVSEKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRRYLAKRQDIKSDGFSIETCKIWMMLDSGSGGLGKEFYIL 600
DB 541 LAGDAEISAFELQTLIRRYLAKRQDIKSDGFSIETCKIWMMLDSGSGGLGKEFYIL 600
QY 601 WTIKIQKQIKYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
DB 601 WTIKIQKQIKYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
QY 661 NFVRCLVRLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCLVRLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 4
US-09-538-092-947
; Sequence 947, Application US/09538092
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; PRIOR APPLICATION NUMBER: 2000-03-29
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

;; NUMBER OF SEQ ID NOS: 1387
;; SOFTWARE: CuraBaseGFormatter Version 0.9
;; SEQ ID NO 947
;; LENGTH: 700
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (0)..(0)
;; OTHER INFORMATION: Polypeptide Accession Number P17655
US-09-538-092-947

Query Match 100.0%; Score 3713; DB 19; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAGIAATLADREABEGLSHERAIKYLNDYELARNCEAGTLFDDPSFPAIPSLGF 60
DB 1 MAGIAATLADREABEGLSHERAIKYLNDYELARNCEAGTLFDDPSFPAIPSLGF 60
QY 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIASTLTNEEI 120
DB 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIASTLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFWQYGEWVEVVDRLPTKGBELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFWQYGEWVEVVDRLPTKGBELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYELKKRPPNLFKIIQKALQKSLG 240
DB 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYELKKRPPNLFKIIQKALQKSLG 240
QY 241 SIITSAADSEALTFOKLVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIITSAADSEALTFOKLVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEBRERLTRRHEDGEFWSFSDFLRHYSRLICNLTPDTLTSPTYKKMLTK 360
DB 301 CPSWNTIDPEBRERLTRRHEDGEFWSFSDFLRHYSRLICNLTPDTLTSPTYKKMLTK 360
QY 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBDEDEBEGSCCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBDEDEBEGSCCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEBELSGQTNHLSKNFLTNARERSDPTFLNREVLNRPFKLP 480
DB 421 KMGEDMHTIGFGIYEVEBELSGQTNHLSKNFLTNARERSDPTFLNREVLNRPFKLP 480
QY 481 EYILVSTFEPNKGDFCIRVSEKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKGDFCIRVSEKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRRYLAKRQDIKSDGFSIETCKIWMMLDSGSGGLGKEFYIL 600
DB 541 LAGDAEISAFELQTLIRRYLAKRQDIKSDGFSIETCKIWMMLDSGSGGLGKEFYIL 600
QY 601 WTIKIQKQIKYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
DB 601 WTIKIQKQIKYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
QY 661 NFVRCLVRLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCLVRLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 5
US-09-562-979-9
; Sequence 9, Application US/09562979
; GENERAL INFORMATION:
; APPLICANT: Fredelky, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS

```
FILE REFERENCE: 24881-301B
CURRENT APPLICATION NUMBER: US/09/562.979
CURRENT FILING DATE: 2000-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-09-562-979-9
```

```
Query Match 100.0%; Score 3713; DB 19; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MAGIAAKLADREABEGLSHERAIKYNODYEALNRECELAAGTLFODPSFPALPSALGF 60
DB 1 MAGIAAKLADREABEGLSHERAIKYNODYEALNRECELAAGTLFODPSFPALPSALGF 60
QY 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWMLAAIASLTINEI 120
DB 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWMLAAIASLTINEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWQGEWVEVVDDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWQGEWVEVVDDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWYELKPPNLFKIIQKALOKGSLG 240
DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWYELKPPNLFKIIQKALOKGSLG 240
QY 241 SIDTSADESEALTFOKLVKGAHAYVTGAEEVNSGSLQKLRIRNPMGEVENTGRMND 300
DB 241 SIDTSADESEALTFOKLVKGAHAYVTGAEEVNSGSLQKLRIRNPMGEVENTGRMND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLTSPTYKKWLT 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLTSPTYKKWLT 360
QY 361 MDGWMRRGSTAGGCRRNYPNTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRROR 420
DB 361 MDGWMRRGSTAGGCRRNYPNTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRROR 420
QY 421 KMGEDMHTIGFIYEVEPEELSGQTNHLSKNFPLTNARERSDPTINREVLNRPKLP 480
DB 421 KMGEDMHTIGFIYEVEPEELSGQTNHLSKNFPLTNARERSDPTINREVLNRPKLP 480
QY 481 EYIIVPSTFEPNKGDFCIRVFSEKADYQAVDDEIEANLEPDISDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVFSEKADYQAVDDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSLETCKIWMMDSDSGSLGKLEFYIL 600
DB 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSLETCKIWMMDSDSGSLGKLEFYIL 600
QY 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHQVIVARFADDOILLIDF 660
DB 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHQVIVARFADDOILLIDF 660
QY 661 NFVRCVLRLETLFKIKQALDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIKQALDPENTGTIELDLISWLCFSVL 700
```

RESULT 6
US-09-768-877-23

```
Sequence 23, Application US/09768877
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ACD:307
CURRENT APPLICATION NUMBER: US/09/768.877
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 09/422,869
PRIOR FILING DATE: 1999-10-21
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-768-877-23
```

```
Query Match 100.0%; Score 3713; DB 22; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MAGIAAKLADREABEGLSHERAIKYNODYEALNRECELAAGTLFODPSFPALPSALGF 60
DB 1 MAGIAAKLADREABEGLSHERAIKYNODYEALNRECELAAGTLFODPSFPALPSALGF 60
QY 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWMLAAIASLTINEI 120
DB 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWMLAAIASLTINEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWQGEWVEVVDDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWQGEWVEVVDDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWYELKPPNLFKIIQKALOKGSLG 240
DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWYELKPPNLFKIIQKALOKGSLG 240
QY 241 SIDTSADESEALTFOKLVKGAHAYVTGAEEVNSGSLQKLRIRNPMGEVENTGRMND 300
DB 241 SIDTSADESEALTFOKLVKGAHAYVTGAEEVNSGSLQKLRIRNPMGEVENTGRMND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLTSPTYKKWLT 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLTSPTYKKWLT 360
QY 361 MDGWMRRGSTAGGCRRNYPNTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRROR 420
DB 361 MDGWMRRGSTAGGCRRNYPNTFMNPOYLKLEBEDDEDESGCTFLVGLIQKRROR 420
QY 421 KMGEDMHTIGFIYEVEPEELSGQTNHLSKNFPLTNARERSDPTINREVLNRPKLP 480
DB 421 KMGEDMHTIGFIYEVEPEELSGQTNHLSKNFPLTNARERSDPTINREVLNRPKLP 480
QY 481 EYIIVPSTFEPNKGDFCIRVFSEKADYQAVDDEIEANLEPDISDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVFSEKADYQAVDDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSLETCKIWMMDSDSGSLGKLEFYIL 600
DB 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSLETCKIWMMDSDSGSLGKLEFYIL 600
QY 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHQVIVARFADDOILLIDF 660
DB 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHQVIVARFADDOILLIDF 660
```

QY 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

RESULT 7

US-09-791-537-81010
; Sequence 81010, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 81010
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-81010

Query Match 100.0%; Score 3713; DB 22; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSGSHERAITYLNODYALNCELEAGTLFQDSPFPAIPALG 60
DB 1 MAGIAKLAKDRBAAGLSGSHERAITYLNODYALNCELEAGTLFQDSPFPAIPALG 60
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKQDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKQDELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWELKPPPLFKIIOKALOKGSLIGC 240
DB 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWELKPPPLFKIIOKALOKGSLIGC 240
QY 241 SIDITSAADSEAITPQKLVKGAHAYVTGAEEVNSGLOKLRIRNPMGEVETGMRNDN 300
DB 241 SIDITSAADSEAITPQKLVKGAHAYVTGAEEVNSGLOKLRIRNPMGEVETGMRNDN 300
QY 301 CPSEMTIDPEERERLTRRHEDGEPFMSFSDFLRHYSRLICNLTPDTLTSDTYKWKLT 360
DB 301 CPSEMTIDPEERERLTRRHEDGEPFMSFSDFLRHYSRLICNLTPDTLTSDTYKWKLT 360
QY 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEDESGCTFLVGLQKHRROR 420
DB 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEDESGCTFLVGLQKHRROR 420
QY 421 KMGEDHTTGFGLYEVPEELSGQTNHLSKNFPLTRARERSDTFNLREVLNRFPLPG 480
DB 421 KMGEDHTTGFGLYEVPEELSGQTNHLSKNFPLTRARERSDTFNLREVLNRFPLPG 480
QY 481 EYILVSTEEPNDGFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYILVSTEEPNDGFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIWDMLDSGSKLGKEFYIL 600
DB 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIWDMLDSGSKLGKEFYIL 600
QY 601 WTKIOYKOKYREIIVDRSGTMSYEMRKALREAGFWMPCQLHOVIVARADOLIIDPD 660
DB 601 WTKIOYKOKYREIIVDRSGTMSYEMRKALREAGFWMPCQLHOVIVARADOLIIDPD 660

QY 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

RESULT 8

US-09-840-707A-9
; Sequence 9, Application US/09840707A
; GENERAL INFORMATION:
; APPLICANT: Fredking, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; FILE REFERENCE: 24881-301C
; CURRENT APPLICATION NUMBER: US/09/840,707A
; CURRENT FILING DATE: 2001-04-23
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/562,979
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: competing amino acid sequence.
; PATENT DOCUMENT NUMBER: 5,817,476
; PUBLICATION DATE: 1995-06-07
US-09-840-707A-9

Query Match 100.0%; Score 3713; DB 23; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSGSHERAITYLNODYALNCELEAGTLFQDSPFPAIPALG 60
DB 1 MAGIAKLAKDRBAAGLSGSHERAITYLNODYALNCELEAGTLFQDSPFPAIPALG 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLAIAISLTINEE 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLAIAISLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKQDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKQDELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWELKPPPLFKIIOKALOKGSLIGC 240
DB 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWELKPPPLFKIIOKALOKGSLIGC 240
QY 241 SIDITSAADSEAITPQKLVKGAHAYVTGAEEVNSGLOKLRIRNPMGEVETGMRNDN 300
DB 241 SIDITSAADSEAITPQKLVKGAHAYVTGAEEVNSGLOKLRIRNPMGEVETGMRNDN 300
QY 301 CPSEMTIDPEERERLTRRHEDGEPFMSFSDFLRHYSRLICNLTPDTLTSDTYKWKLT 360
DB 301 CPSEMTIDPEERERLTRRHEDGEPFMSFSDFLRHYSRLICNLTPDTLTSDTYKWKLT 360
QY 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEDESGCTFLVGLQKHRROR 420
DB 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEDESGCTFLVGLQKHRROR 420
QY 421 KMGEDHTTGFGLYEVPEELSGQTNHLSKNFPLTRARERSDTFNLREVLNRFPLPG 480
DB 421 KMGEDHTTGFGLYEVPEELSGQTNHLSKNFPLTRARERSDTFNLREVLNRFPLPG 480
QY 481 EYILVSTEEPNDGFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYILVSTEEPNDGFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDGVRRLFAQ 540

Db 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDIDEDDGVRLFAQ 540
QY 541 LAGEDAISAPELQTLIRVLAKRQDIKSDGFSIETCKIMVMDSDSGKGLKEFYIL 600
Db 541 LAGEDAISAPELQTLIRVLAKRQDIKSDGFSIETCKIMVMDSDSGKGLKEFYIL 600
QY 601 WTKIQYKQYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
Db 601 WTKIQYKQYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
Db 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700

RESULT 9

US-09-840-707B-9
; Sequence 9, Application US/09840707B
; GENERAL INFORMATION:
; APPLICANT: Fredeking, Terry M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
; FILE REFERENCE: 24881-301C
; CURRENT APPLICATION NUMBER: US/09/840,707B
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/562,979
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/198,210
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: comprising amino acid sequence
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5,817,476
; PATENT FILING DATE: 1995-06-07
; PUBLICATION DATE: 1998-10-06
; US-09-840-707B-9

Query Match 100.0%; Score 3713; DB 23; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKREAAEGLSHERAIKYNQVEALRNECLBAGTLFQDPSPALPSALGF 60
Db 1 MAGIAAKLAKREAAEGLSHERAIKYNQVEALRNECLBAGTLFQDPSPALPSALGF 60
QY 61 KELGPYSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWLLAIASTLTNEI 120
Db 61 KELGPYSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWLLAIASTLTNEI 120
QY 121 LARVVPNOSFOENAGIFHFQFQWYGEWVEVVDRLPTKDGELLFVHSAGSEFWAL 180
Db 121 LARVVPNOSFOENAGIFHFQFQWYGEWVEVVDRLPTKDGELLFVHSAGSEFWAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEFGIAEWELKKPPNLFKIQKLOKSLIGC 240
Db 181 LEKAVAKINGCYEALSGATTEGFEFGIAEWELKKPPNLFKIQKLOKSLIGC 240
QY 241 SIDITSADSEALITFQKLVKHAIVTGAEEVNSGSLQKLRIRNPGVEVWTRMND 300
Db 241 SIDITSADSEALITFQKLVKHAIVTGAEEVNSGSLQKLRIRNPGVEVWTRMND 300
QY 301 CPSWNTIDPERERLTRHEDGEEFWMSDFLRHYSRLICNLTPDTLTSDTYKKMKLT 360
Db 301 CPSWNTIDPERERLTRHEDGEEFWMSDFLRHYSRLICNLTPDTLTSDTYKKMKLT 360

QY 361 MDGNMRGSGTAGGCNRYNPTFMNNPOYLILKEBEDDEEDSGSGCTFLVGLIQKRRROR 420
Db 361 MDGNMRGSGTAGGCNRYNPTFMNNPOYLILKEBEDDEEDSGSGCTFLVGLIQKRRROR 420
QY 421 KMEDNHTTGFQYVEPEELSGQTNHLKSNFELTRARERSTFTNLREVNREFLPG 480
Db 421 KMEDNHTTGFQYVEPEELSGQTNHLKSNFELTRARERSTFTNLREVNREFLPG 480
QY 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDIDEDDGVRLFAQ 540
Db 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDIDEDDGVRLFAQ 540
QY 541 LAGEDAISAPELQTLIRVLAKRQDIKSDGFSIETCKIMVMDSDSGKGLKEFYIL 600
Db 541 LAGEDAISAPELQTLIRVLAKRQDIKSDGFSIETCKIMVMDSDSGKGLKEFYIL 600
QY 601 WTKIQYKQYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
Db 601 WTKIQYKQYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
Db 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700

RESULT 10

US-09-884-319-7
; Sequence 7, Application US/09884319
; GENERAL INFORMATION:
; APPLICANT: Lin, Lih-Ling
; TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
; INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
; BINDING

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,319
FILING DATE: 18-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,516
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/487,942
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-884-319-7

Query Match 100.0%; Score 3713; DB 23; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGIAKLADRAAAGLGSHERAIKYNODYALNECELEAGTLFODSPFPAIPALGF 60
 1 MAGIAKLADRAAAGLGSHERAIKYNODYALNECELEAGTLFODSPFPAIPALGF 60
 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCMLAAIASLTINEE 120
 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCMLAAIASLTINEE 120
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKOGELLFVHSAEGSEFMSAL 180
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKOGELLFVHSAEGSEFMSAL 180
 181 LEKAYAKINGCYALSGGATTBGEFDTGIAEWYELKKPPNLFKTIQKALOKSILGC 240
 181 LEKAYAKINGCYALSGGATTBGEFDTGIAEWYELKKPPNLFKTIQKALOKSILGC 240
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEVESNSLOKLRIRNPMGEVETGRWNND 300
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEVESNSLOKLRIRNPMGEVETGRWNND 300
 301 CPSMNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 301 CPSMNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 361 MDGNMRRGSTAGGCNRYPTFMNPOYLKLEEDDEDEBEGSGCTFVLGLIOKRRROR 420
 361 MDGNMRRGSTAGGCNRYPTFMNPOYLKLEEDDEDEBEGSGCTFVLGLIOKRRROR 420
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLNKFPLTRARERSPTFLNREVLNRPFLPG 480
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLNKFPLTRARERSPTFLNREVLNRPFLPG 480
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDEIEANLEBDEIDEDIDGVRRLFAQ 540
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDEIEANLEBDEIDEDIDGVRRLFAQ 540
 541 LAGEADAISAFELQTLIRVLAKRDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
 541 LAGEADAISAFELQTLIRVLAKRDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
 601 WTKIOYOKIYRIIDVDSGTMSYEMRKALEAGFMPQLHOVIVARPADQLIIDFD 660
 601 WTKIOYOKIYRIIDVDSGTMSYEMRKALEAGFMPQLHOVIVARPADQLIIDFD 660
 661 NFVRCVLRLETLFKIFKQLDPEMTGTELDLISMLCPSVL 700
 661 NFVRCVLRLETLFKIFKQLDPEMTGTELDLISMLCPSVL 700

RESULT 11

US-09-884-319A-7

Sequence 7, Application US/09884319A

GENERAL INFORMATION:

APPLICANT: Lit, Lih-Ling

ADDRESS: Graham, James

TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR

INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND

BINDING

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: USA

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/884,319A

FILING DATE: 18-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/083,516

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/487,942

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

REFERENCE/DOCKET NUMBER: G15258

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 700 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULAR TYPE: protein

HYPOTHETICAL: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 100.0%; Score 3713; DB 23; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGIAKLADRAAAGLGSHERAIKYNODYALNECELEAGTLFODSPFPAIPALGF 60
 1 MAGIAKLADRAAAGLGSHERAIKYNODYALNECELEAGTLFODSPFPAIPALGF 60
 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCMLAAIASLTINEE 120
 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCMLAAIASLTINEE 120
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKOGELLFVHSAEGSEFMSAL 180
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKOGELLFVHSAEGSEFMSAL 180
 181 LEKAYAKINGCYALSGGATTBGEFDTGIAEWYELKKPPNLFKTIQKALOKSILGC 240
 181 LEKAYAKINGCYALSGGATTBGEFDTGIAEWYELKKPPNLFKTIQKALOKSILGC 240
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEVESNSLOKLRIRNPMGEVETGRWNND 300
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEVESNSLOKLRIRNPMGEVETGRWNND 300
 301 CPSMNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 301 CPSMNTIDPEERERLTRRHEDGFPMWSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 361 MDGNMRRGSTAGGCNRYPTFMNPOYLKLEEDDEDEBEGSGCTFVLGLIOKRRROR 420
 361 MDGNMRRGSTAGGCNRYPTFMNPOYLKLEEDDEDEBEGSGCTFVLGLIOKRRROR 420
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLNKFPLTRARERSPTFLNREVLNRPFLPG 480
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLNKFPLTRARERSPTFLNREVLNRPFLPG 480
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDEIEANLEBDEIDEDIDGVRRLFAQ 540
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDEIEANLEBDEIDEDIDGVRRLFAQ 540
 541 LAGEADAISAFELQTLIRVLAKRDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
 541 LAGEADAISAFELQTLIRVLAKRDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600

QY 601 WTGIQYKTYREIDVDSGTMSYEMRKALBEAGFMPQCHQVIVARFADQLIIDFD 660
DB 601 WTGIQYKTYREIDVDSGTMSYEMRKALBEAGFMPQCHQVIVARFADQLIIDFD 660
QY 661 NFVRCVRLTETLFKIFKQDPDENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDPDENTGTIELDLISWLCFSVL 700

RESULT 12

US-10-038-557A-9
Sequence 9, Application US/10038557A
GENERAL INFORMATION:
APPLICANT: Fredeking, Terry M.
APPLICANT: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
FILE REFERENCE: 24881-301D
CURRENT APPLICATION NUMBER: US/10/038, 557A
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 09/840,707
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/562,979
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/198,210
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
OTHER INFORMATION: comprising amino acid sequence
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-10-038-557A-9

Query Match 100.0%; Score 3713; DB 26; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKDREAELGSGHERAIKYLNDYALNECEAGTLPDPSFPALPSALGF 60
DB 1 MAGIAAKLAKDREAELGSGHERAIKYLNDYALNECEAGTLPDPSFPALPSALGF 60
QY 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAIATLTNEEI 120
DB 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAIATLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
QY 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
DB 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
QY 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
DB 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
QY 241 SIDITSAADSEATTFQKLVKGHAYVTGAEEVESNGSLQKILIRNPMGEVEMTGRMNDN 300
DB 241 SIDITSAADSEATTFQKLVKGHAYVTGAEEVESNGSLQKILIRNPMGEVEMTGRMNDN 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
QY 361 MDGWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEBDESGCTPLVGLIQHRRROR 420
DB 361 MDGWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEBDESGCTPLVGLIQHRRROR 420

QY 421 KMGEDNHTIGFGIYEVPEELSGQTNILSKNPELJINRAERSDTFINLEVINRFXLPFG 480
DB 421 KMGEDNHTIGFGIYEVPEELSGQTNILSKNPELJINRAERSDTFINLEVINRFXLPFG 480
QY 481 EYTLVSTPEPNKDGPCFIRVSEKADYQAVNDELEAULEEEDIEDDIDGVRRLFAQ 540
DB 481 EYTLVSTPEPNKDGPCFIRVSEKADYQAVNDELEAULEEEDIEDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTLIRVLAARODIKSGFSIETCKIMVDLSDSGSKGLKEFPYL 600
DB 541 LAGEDAEISAFELQTLIRVLAARODIKSGFSIETCKIMVDLSDSGSKGLKEFPYL 600
QY 601 WTGIQYKTYREIDVDSGTMSYEMRKALBEAGFMPQCHQVIVARFADQLIIDFD 660
DB 601 WTGIQYKTYREIDVDSGTMSYEMRKALBEAGFMPQCHQVIVARFADQLIIDFD 660
QY 661 NFVRCVRLTETLFKIFKQDPDENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDPDENTGTIELDLISWLCFSVL 700

RESULT 13

US-10-116-519-12
Sequence 12, Application US/10116519
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
TITLE OF INVENTION: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
FILE REFERENCE: D0124 NP
CURRENT APPLICATION NUMBER: US/10/116, 519
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/281,253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288,768
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/296,180
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/300,620
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
US-10-116-519-12

Query Match 100.0%; Score 3713; DB 27; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKDREAELGSGHERAIKYLNDYALNECEAGTLPDPSFPALPSALGF 60
DB 1 MAGIAAKLAKDREAELGSGHERAIKYLNDYALNECEAGTLPDPSFPALPSALGF 60
QY 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAIATLTNEEI 120
DB 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAIATLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVDDRLPTKDELLFVHSAEGSEFWSAL 180
QY 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
DB 181 LEKAYAKINGCYALSGATTEGFEEDFTGGIAEWYELKKPPNLFKIIQKALOKGSLGC 240
QY 241 SIDITSAADSEATTFQKLVKGHAYVTGAEEVESNGSLQKILIRNPMGEVEMTGRMNDN 300
DB 241 SIDITSAADSEATTFQKLVKGHAYVTGAEEVESNGSLQKILIRNPMGEVEMTGRMNDN 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360

Db 301 CPSMNTIDPEERERLTRRHEDGFFWMSFSDPLRHYSRLEICNLTPTDLSDTYKKWLTk 360
Qy 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEBEGSGCTFLVGLIQKHRROR 420
Db 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEBEGSGCTFLVGLIQKHRROR 420
Qy 421 KMGEDMHTTGFGLIYEPBEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Db 421 KMGEDMHTTGFGLIYEPBEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Qy 481 EYLLVSTFEPNKGDFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDDGVRRLFAQ 540
Db 481 EYLLVSTFEPNKGDFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDDGVRRLFAQ 540
Qy 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKFEYIL 600
Db 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKFEYIL 600
Qy 601 WTKIQYQKTYREIDVDRSGTMSYEMRKALBEAGFMPCOLHQVIVARPADQOLIIFD 660
Db 601 WTKIQYQKTYREIDVDRSGTMSYEMRKALBEAGFMPCOLHQVIVARPADQOLIIFD 660
Qy 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700

RESULT 14

US-10-219-051B-12485
; Sequence 12485, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 12485
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Refseq / NP_001739
; DATABASE ENTRY DATE: 2001-01-23
US-10-219-051B-12485

Query Match 100.0%; Score 3713; DB 28; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIAKLAKDRRAAGLSGSHERAIKYLNQDYBALNECLEAGTLFQDPSFPAIPALGP 60
Db 1 MAGIAKLAKDRRAAGLSGSHERAIKYLNQDYBALNECLEAGTLFQDPSFPAIPALGP 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTINEE 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTINEE 120
Qy 121 LARVVLNOSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAESSEFWSAL 180
Db 121 LARVVLNOSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAESSEFWSAL 180
Qy 181 LERAYAKINCCEALSGATTEGDEDTGGIAEWELKRPPLFKIQALOKGSLIGC 240
Db 181 LERAYAKINCCEALSGATTEGDEDTGGIAEWELKRPPLFKIQALOKGSLIGC 240

Db 181 LERAYAKINCCEALSGATTEGDEDTGGIAEWELKRPPLFKIQALOKGSLIGC 240
Qy 241 SIDITSAADSEAITPOKLVKHAYSVTGAEEVENSGLQKLIIRNPMGVEWETGRMND 300
Db 241 SIDITSAADSEAITPOKLVKHAYSVTGAEEVENSGLQKLIIRNPMGVEWETGRMND 300
Qy 301 CPSMNTIDPEERERLTRRHEDGFFWMSFSDPLRHYSRLEICNLTPTDLSDTYKKWLTk 360
Db 301 CPSMNTIDPEERERLTRRHEDGFFWMSFSDPLRHYSRLEICNLTPTDLSDTYKKWLTk 360
Qy 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEBEGSGCTFLVGLIQKHRROR 420
Db 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEBEGSGCTFLVGLIQKHRROR 420
Qy 421 KMGEDMHTTGFGLIYEPBEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Db 421 KMGEDMHTTGFGLIYEPBEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Qy 481 EYLLVSTFEPNKGDFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDDGVRRLFAQ 540
Db 481 EYLLVSTFEPNKGDFCIRVFSEKKADYQAVDEIEANLEEFDISDDIDDGVRRLFAQ 540
Qy 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKFEYIL 600
Db 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKFEYIL 600
Qy 601 WTKIQYQKTYREIDVDRSGTMSYEMRKALBEAGFMPCOLHQVIVARPADQOLIIFD 660
Db 601 WTKIQYQKTYREIDVDRSGTMSYEMRKALBEAGFMPCOLHQVIVARPADQOLIIFD 660
Qy 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700

RESULT 15

US-10-293-017-78
; Sequence 78, Application US/10293017
; GENERAL INFORMATION:
; APPLICANT: Shamloo, Mehrdad
; APPLICANT: Chin, Daniel J.
; APPLICANT: Von Schack, David
; APPLICANT: Ufer, Roman
; APPLICANT: Palfy, Laszlo
; APPLICANT: Wieloch, Tadeusz
; APPLICANT: Gonzalez-Zulueta, Mirella
; TITLE OF INVENTION: Regulated Genes in the Pathophysiology of Ischemic Stroke
; FILE REFERENCE: AGYT-007US1
; CURRENT APPLICATION NUMBER: US/10/293,017
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 60/384,522
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/338,296
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-017-78

Query Match 100.0%; Score 3713; DB 28; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIAKLAKDRRAAGLSGSHERAIKYLNQDYBALNECLEAGTLFQDPSFPAIPALGP 60
Db 1 MAGIAKLAKDRRAAGLSGSHERAIKYLNQDYBALNECLEAGTLFQDPSFPAIPALGP 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTINEE 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTINEE 120

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:53:55 ; Search time 47.4084 Seconds
(without alignments)
1753.524 Million cell updates/sec

Title: US-09-884-319a-7
Perfect score: 3713
Sequence: 1 MAGIAKLADREAAGLSGSGS.....PENTGTEILDLSMLCPSTVL 700

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCRT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCRTUS_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	US-09-840-707A-9	Sequence 9, Appl1
2	3713	100.0	700	US-09-768-877-23	Sequence 23, Appl1
3	3713	100.0	700	US-09-884-319-7	Sequence 7, Appl1
4	3713	100.0	700	US-10-038-557A-9	Sequence 9, Appl1
5	3713	100.0	700	US-10-116-519-12	Sequence 12, Appl1
6	2458.5	66.2	714	US-09-768-877-22	Sequence 22, Appl1
7	2458.5	66.2	714	US-10-116-519-6	Sequence 6, Appl1
8	2435	65.6	713	US-10-116-519-8	Sequence 8, Appl1
9	2356.5	63.5	702	US-09-768-877-27	Sequence 27, Appl1
10	2014	54.2	703	US-10-116-519-11	Sequence 11, Appl1
11	1943.5	52.3	709	US-10-116-519-9	Sequence 9, Appl1
12	1882	50.7	690	US-09-768-877-28	Sequence 28, Appl1
13	1882	50.7	690	US-10-116-519-13	Sequence 13, Appl1
14	1862.5	50.2	821	US-09-768-877-24	Sequence 24, Appl1
15	1862.5	50.2	821	US-10-116-519-7	Sequence 7, Appl1

16	1795	48.3	664	US-10-325-881-70	Sequence 70, Appl1
17	1783.5	48.0	647	US-10-116-519-10	Sequence 10, Appl1
18	1287	34.7	251	US-09-840-707A-8	Sequence 8, Appl1
19	1287	34.7	251	US-09-884-319-6	Sequence 6, Appl1
20	1287	34.7	251	US-10-038-557A-8	Sequence 8, Appl1
21	1150	31.0	697	US-10-116-519-56	Sequence 56, Appl1
22	1148.5	30.9	694	US-10-116-519-54	Sequence 54, Appl1
23	1126.5	30.3	343	US-09-794-960-4	Sequence 4, Appl1
24	1056	28.4	206	US-10-043-487-221	Sequence 221, Appl1
25	1020	27.5	300	US-09-794-960-5	Sequence 5, Appl1
26	938	25.3	428	US-10-116-519-2	Sequence 2, Appl1
27	938	25.3	428	US-10-116-519-25	Sequence 25, Appl1
28	901.5	24.3	639	US-09-768-877-25	Sequence 25, Appl1
29	893	24.1	639	US-10-116-519-5	Sequence 5, Appl1
30	845.5	22.8	669	US-09-962-739-2	Sequence 2, Appl1
31	780	21.0	641	US-09-768-877-26	Sequence 26, Appl1
32	656	17.7	353	US-09-962-739-4	Sequence 4, Appl1
33	635	17.1	242	US-10-116-519-92	Sequence 92, Appl1
34	635	17.1	242	US-10-116-519-93	Sequence 93, Appl1
35	632.5	17.0	241	US-10-116-519-46	Sequence 46, Appl1
36	621.5	16.7	666	US-09-768-877-18	Sequence 18, Appl1
37	607	16.3	544	US-09-768-877-4	Sequence 4, Appl1
38	604	16.3	671	US-10-116-519-4	Sequence 4, Appl1
39	604	16.3	672	US-09-768-877-2	Sequence 2, Appl1
40	600	16.2	1603	US-10-208-948-2	Sequence 2, Appl1
41	600	16.2	2159	US-10-208-948-24	Sequence 24, Appl1
42	599	16.1	513	US-09-768-877-8	Sequence 8, Appl1
43	579	15.6	517	US-09-794-960-2	Sequence 2, Appl1
44	575	15.5	517	US-09-768-877-6	Sequence 6, Appl1
45	561	15.1	2150	US-10-208-948-27	Sequence 27, Appl1

ALIGNMENTS

RESULT 1
US-09-840-707A-9
Sequence 9, Application US/09840707A
Patent No. US20020077276A1
GENERAL INFORMATION:
APPLICANT: Predelkay, Terry M.
INVENTOR: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
FILE REFERENCE: 24881-301C
CURRENT APPLICATION NUMBER: US/09/840, 707A
CURRENT FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/562, 979
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/198, 210
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5, 817, 476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-09-840-707A-9

Query Match 100.0%; Score 3713; DB 9; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 MAGIAKLADREAAGLSGSGSHERAIKYLNDYALRNECLEAGTLFDDPSFPAISALGF 60
DB 1 MAGIAKLADREAAGLSGSGSHERAIKYLNDYALRNECLEAGTLFDDPSFPAISALGF 60

```

QY      61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTINEI 120
        61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTINEI 120
Db      121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
        121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
QY      181 LERAYAKINGCYEALSGGATTEGFEDFTGIAEMWELKKEPPNLFKIIQKALOKGSLIGC 240
        181 LERAYAKINGCYEALSGGATTEGFEDFTGIAEMWELKKEPPNLFKIIQKALOKGSLIGC 240
Db      241 SIDTSADESEAITFOKLVGHAYSVTGAEEVENSGLQKLIIRINPWGEVMTGRWMDN 300
        241 SIDTSADESEAITFOKLVGHAYSVTGAEEVENSGLQKLIIRINPWGEVMTGRWMDN 300
QY      301 CPSWNTIDPEERERLTRRHDEGFWMFSFDFLRHYSRLIECNLTPTDITSDTYKKMKLTJ 360
        301 CPSWNTIDPEERERLTRRHDEGFWMFSFDFLRHYSRLIECNLTPTDITSDTYKKMKLTJ 360
Db      361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
        361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
QY      421 KMGEDMTTIGFGIYEVEPEELSGQTNILSKNFFLTNRARSDPTFNLREVLRNFKLPJG 480
        421 KMGEDMTTIGFGIYEVEPEELSGQTNILSKNFFLTNRARSDPTFNLREVLRNFKLPJG 480
Db      481 EYIIVPSTFEPNKGDFCIRVFSEKKADYQAVDDEIEANLEEDISDDIDDGVRRLFAQ 540
        481 EYIIVPSTFEPNKGDFCIRVFSEKKADYQAVDDEIEANLEEDISDDIDDGVRRLFAQ 540
QY      541 LAGDAEISAFELQTLIRRYLAKRODICKSGFSIETCKIWMMLDSGSKGLKEFYIL 600
        541 LAGDAEISAFELQTLIRRYLAKRODICKSGFSIETCKIWMMLDSGSKGLKEFYIL 600
Db      601 WTKIQKQKXIREIDVDRSGTMSYEMRKALBEGAFKPCQLHQVYARFADQLIIDFD 660
        601 WTKIQKQKXIREIDVDRSGTMSYEMRKALBEGAFKPCQLHQVYARFADQLIIDFD 660
QY      661 NFVRCLVRLTFLFKIFKQDLPENTGTIELDLISWLCFSVL 700
        661 NFVRCLVRLTFLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 2
US-09-768-877-23

```

; Sequence 23, Application US/09768877
; Patent No. US20020150896A1
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/768, 877
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/422, 869
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 700
; TYPE: PRF
; ORGANISM: Human
US-09-768-877-23

```

Query Match 100.0%; Score 3713; DB 10; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MAGIAAALADREABEGSHERAIKYNODYALRNECELAGTLPDPSFPAPISALGF 60
        1 MAGIAAALADREABEGSHERAIKYNODYALRNECELAGTLPDPSFPAPISALGF 60
Db      61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTINEI 120
        61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTINEI 120
QY      121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
        121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
Db      181 LERAYAKINGCYEALSGGATTEGFEDFTGIAEMWELKKEPPNLFKIIQKALOKGSLIGC 240
        181 LERAYAKINGCYEALSGGATTEGFEDFTGIAEMWELKKEPPNLFKIIQKALOKGSLIGC 240
QY      241 SIDTSADESEAITFOKLVGHAYSVTGAEEVENSGLQKLIIRINPWGEVMTGRWMDN 300
        241 SIDTSADESEAITFOKLVGHAYSVTGAEEVENSGLQKLIIRINPWGEVMTGRWMDN 300
Db      301 CPSWNTIDPEERERLTRRHDEGFWMFSFDFLRHYSRLIECNLTPTDITSDTYKKMKLTJ 360
        301 CPSWNTIDPEERERLTRRHDEGFWMFSFDFLRHYSRLIECNLTPTDITSDTYKKMKLTJ 360
QY      361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
        361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
Db      421 KMGEDMTTIGFGIYEVEPEELSGQTNILSKNFFLTNRARSDPTFNLREVLRNFKLPJG 480
        421 KMGEDMTTIGFGIYEVEPEELSGQTNILSKNFFLTNRARSDPTFNLREVLRNFKLPJG 480
QY      481 EYIIVPSTFEPNKGDFCIRVFSEKKADYQAVDDEIEANLEEDISDDIDDGVRRLFAQ 540
        481 EYIIVPSTFEPNKGDFCIRVFSEKKADYQAVDDEIEANLEEDISDDIDDGVRRLFAQ 540
Db      541 LAGDAEISAFELQTLIRRYLAKRODICKSGFSIETCKIWMMLDSGSKGLKEFYIL 600
        541 LAGDAEISAFELQTLIRRYLAKRODICKSGFSIETCKIWMMLDSGSKGLKEFYIL 600
QY      601 WTKIQKQKXIREIDVDRSGTMSYEMRKALBEGAFKPCQLHQVYARFADQLIIDFD 660
        601 WTKIQKQKXIREIDVDRSGTMSYEMRKALBEGAFKPCQLHQVYARFADQLIIDFD 660
QY      661 NFVRCLVRLTFLFKIFKQDLPENTGTIELDLISWLCFSVL 700
        661 NFVRCLVRLTFLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 3
US-09-884-319-7

```

; Sequence 7, Application US/09884319
; Publication No. US20030124625A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Lih-ling
; APPLICANT: Graham, James
; TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
; INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
; BINDING
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,319
FILING DATE: 18-Jun-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/083,516
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/487,942
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-884-319-7

Query Match 100.0%; Score 3713; DB 11; Length 700;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSGHERAIKYNODYALNNECLAEAGTLFODPSFPAIPSAAGF 60
DB 1 MAGIAKLAKDRBAAGLSGHERAIKYNODYALNNECLAEAGTLFODPSFPAIPSAAGF 60
QY 61 KELGPYSSKTRGRMRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTNNEI 120
DB 61 KELGPYSSKTRGRMRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTNNEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYALSGGATTBGFEDFTGGIAEWELKKPPNLFKIQALOKGSLGC 240
DB 181 LERAYAKINGCYALSGGATTBGFEDFTGGIAEWELKKPPNLFKIQALOKGSLGC 240
QY 241 SIDITSAADSEALTFOKLKYGHAYSVTGAEEVENSLSLOKLRIRNPWGEVETGRWNN 300
DB 241 SIDITSAADSEALTFOKLKYGHAYSVTGAEEVENSLSLOKLRIRNPWGEVETGRWNN 300
QY 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
DB 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
QY 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
DB 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
QY 361 MDGNMRRGSTAGGCRRNPNTFMNPOYLKLEEBEDEDEBDESGCTFLVGLIQHRRROR 420
DB 361 MDGNMRRGSTAGGCRRNPNTFMNPOYLKLEEBEDEDEBDESGCTFLVGLIQHRRROR 420
QY 421 KMGEDHTTIFGFIYEVEBELSGOTNHLKSNFPLTNARERSDPTFNLRLVLRPFKLPFG 480
DB 421 KMGEDHTTIFGFIYEVEBELSGOTNHLKSNFPLTNARERSDPTFNLRLVLRPFKLPFG 480
QY 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEEDISEDDIDDGVRLLFAQ 540
DB 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEEDISEDDIDDGVRLLFAQ 540
QY 541 LAEDDAISAFELQTIIRYLAKQDIKSGFSIETTKIIVMDLDSGSGKLKKEFYIL 600
DB 541 LAEDDAISAFELQTIIRYLAKQDIKSGFSIETTKIIVMDLDSGSGKLKKEFYIL 600

DB 541 LAEDDAISAFELQTIIRYLAKQDIKSGFSIETTKIIVMDLDSGSGKLKKEFYIL 600
QY 601 WTIOKQYKQIYREIDVDSGTMSYMRKALAEAGFKMPQLHQVIVAREADQDILDFD 660
DB 601 WTIOKQYKQIYREIDVDSGTMSYMRKALAEAGFKMPQLHQVIVAREADQDILDFD 660
QY 661 NFVRCVLRLETLFKIFKQDIPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQDIPENTGTIELDLISWLCFSVL 700

RESULT 4

US-10-038-557A-9
Sequence 9, Application US/10038557A
Publication No. US20030092684A1
GENERAL INFORMATION:
APPLICANT: Fredeking, Terry M.
APPLICANT: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
FILE REFERENCE: 24861-301D
CURRENT APPLICATION NUMBER: US/10/038,557A
PRIORITY FILING DATE: 2002-06-05
PRIORITY FILING DATE: 09/840,707
PRIORITY FILING DATE: 2001-04-23
PRIORITY FILING DATE: 09/562,979
PRIORITY FILING DATE: 2000-04-27
PRIORITY FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-10-038-557A-9

Query Match 100.0%; Score 3713; DB 15; Length 700;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSGHERAIKYNODYALNNECLAEAGTLFODPSFPAIPSAAGF 60
DB 1 MAGIAKLAKDRBAAGLSGHERAIKYNODYALNNECLAEAGTLFODPSFPAIPSAAGF 60
QY 61 KELGPYSSKTRGRMRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTNNEI 120
DB 61 KELGPYSSKTRGRMRPTEICADPOFIIGATRTDICGALGDCWLLAIAISLTNNEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYALSGGATTBGFEDFTGGIAEWELKKPPNLFKIQALOKGSLGC 240
DB 181 LERAYAKINGCYALSGGATTBGFEDFTGGIAEWELKKPPNLFKIQALOKGSLGC 240
QY 241 SIDITSAADSEALTFOKLKYGHAYSVTGAEEVENSLSLOKLRIRNPWGEVETGRWNN 300
DB 241 SIDITSAADSEALTFOKLKYGHAYSVTGAEEVENSLSLOKLRIRNPWGEVETGRWNN 300
QY 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
DB 301 CPSPMTIDPEERERLRRHEDGEFWSFSDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
QY 361 MDGNMRRGSTAGGCRRNPNTFMNPOYLKLEEBEDEDEBDESGCTFLVGLIQHRRROR 420
DB 361 MDGNMRRGSTAGGCRRNPNTFMNPOYLKLEEBEDEDEBDESGCTFLVGLIQHRRROR 420

```
Db 361 MDGNWRGSGTAGCCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMTTIGFGIYEVEPEELSGOTNHLKSNFELTNRARSDPTFNLREVLNFKLPKPG 480
Db 421 KMGEDMTTIGFGIYEVEPEELSGOTNHLKSNFELTNRARSDPTFNLREVLNFKLPKPG 480
Qy 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Db 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIEISAFELQTLIRRLVAKRQDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
Db 541 LAGEDAIEISAFELQTLIRRLVAKRQDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
Qy 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPQQLHOVYVARPADQLIIDFD 660
Db 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPQQLHOVYVARPADQLIIDFD 660
Qy 661 NFVRCVLRLETLFKIFKQDLPENTGTIELDLISMLCFSVL 700
Db 661 NFVRCVLRLETLFKIFKQDLPENTGTIELDLISMLCFSVL 700
```

```
RESULT 5
US-10-116-519-12
; Sequence 12, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-519-12
```

```
Query Match 100.0%; Score 3713; DB 15; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIAAKLAKDREAEGLGSHERAIKYLNQDYEAALRNECLBAGTLFQDPSFPALPSALGF 60
Db 1 MAGIAAKLAKDREAEGLGSHERAIKYLNQDYEAALRNECLBAGTLFQDPSFPALPSALGF 60
Qy 61 KELGPISSKTRGMRMKRPTTEICADPQITIGATRTDICOAGLGDCLAAIASLTINEEI 120
Db 61 KELGPISSKTRGMRMKRPTTEICADPQITIGATRTDICOAGLGDCLAAIASLTINEEI 120
Qy 121 LARVVPINQSFQENYAGIFHFQFMQYGEWVAVDDRLPTDQGLLFFVHSAEGSEFWSAL 180
Db 121 LARVVPINQSFQENYAGIFHFQFMQYGEWVAVDDRLPTDQGLLFFVHSAEGSEFWSAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDEFGIAEWTELKPPNLFKTIQKLOKSGSLIGC 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDEFGIAEWTELKPPNLFKTIQKLOKSGSLIGC 240
Qy 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEVESNGSLQKIRIRNPGVEWNTGRWMDN 300
Db 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEVESNGSLQKIRIRNPGVEWNTGRWMDN 300
```

```
Db 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEVESNGSLQKIRIRNPGVEWNTGRWMDN 300
Qy 301 CPSMNTIDPEERERLIRRHEDGEFMMSPDFLHYRLEICNLTPDLTSDTYKKWLTJK 360
Db 301 CPSMNTIDPEERERLIRRHEDGEFMMSPDFLHYRLEICNLTPDLTSDTYKKWLTJK 360
Qy 361 MDGNWRGSGTAGCCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Db 361 MDGNWRGSGTAGCCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMTTIGFGIYEVEPEELSGOTNHLKSNFELTNRARSDPTFNLREVLNFKLPKPG 480
Db 421 KMGEDMTTIGFGIYEVEPEELSGOTNHLKSNFELTNRARSDPTFNLREVLNFKLPKPG 480
Qy 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Db 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIEISAFELQTLIRRLVAKRQDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
Db 541 LAGEDAIEISAFELQTLIRRLVAKRQDIKSDGFSIETCKIMVMDLSDSGSKLGKEFYIL 600
Qy 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPQQLHOVYVARPADQLIIDFD 660
Db 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPQQLHOVYVARPADQLIIDFD 660
Qy 661 NFVRCVLRLETLFKIFKQDLPENTGTIELDLISMLCFSVL 700
Db 661 NFVRCVLRLETLFKIFKQDLPENTGTIELDLISMLCFSVL 700
```

```
RESULT 6
US-09-768-877-22
; Sequence 22, Application US/09768877
; Patent No. US20020150896A1
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/768,877
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/422,869
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Human
US-09-768-877-22
```

```
Query Match 66.2%; Score 2458.5; DB 10; Length 714;
Best Local Similarity 63.1%; Pred. No. 6,3e-221;
Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

Qy 3 GIAAKLAKDREAEGLGSHERAIKYLNQDYEAALRNECLBAGTLFQDPSFPALPSALGF 62
Db 3 GVAHQVQKQARELIGLRHENAIAIKYLGQDYEQLRVRLQSGTLFRDAEPVPOQSLGYKO 72
Qy 73 LGRNRSKTYIGIKMRPPELLSNQFIVDGAIRTDICOAGLGDCLAAIASLTINDTLH 132
Db 73 LGRNRSKTYIGIKMRPPELLSNQFIVDGAIRTDICOAGLGDCLAAIASLTINDTLH 132
Qy 123 RVPPLNQSFOENYAGIFHFQFMQYGEWVAVDDRLPTDQGLLFFVHSAEGSEFWSAL 182
Db 123 RVPPLNQSFOENYAGIFHFQFMQYGEWVAVDDRLPTDQGLLFFVHSAEGSEFWSAL 182
```

```

Db      133 RVPFHGSGFONGAGIHFHOLMOGEWVDVVDLLPIKDGKLVFHSAGNEFWMSALLE 192
Qy      183 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIOKALOKSILGCSI 242
      193 KAVAKONGSYEALSGSTSGFEDFTGIVTEWELKRAPSDLYOIILKALERSILGCSI 252
Qy      243 DITSADSEAITFOKLVKHAIVTGAEEVNSGLOKLRIRINPGEVMTGRANDNCP 302
      253 DISSVADMEAITFKLVKHAIVTGAQVNVGQVSLIRMRNPMGEVMTGAMSDSS 312
Qy      303 SMTIDPEERERLTRRHEDGEFWMSPDLRHYSRLICNLTPDTLSDTYKKKLTMD 362
      313 EMNVVDYERDQLRVKMEDEFPWMSFRDPFRBEFTRLICNLTPDALKSRTIRKNTTLYE 372
Qy      363 GNMRRGSTAGGCNYPNTFMMNPQVYLKLEEDDEE--DGESGCTFVLGLIOKRRROR 420
      373 GTMRSGSTAGGCNYPATFVWNPQFKRLDETDPDYGDRSGCSFVLALMQHRRER 432
Qy      421 KMGEDMHTTGFYIEVPEELSGQTNHLSKNFPLTNRAERSDTFNLREVLNRFKLP 480
      433 RFRDMETTGFAYEVPPELVGQPAVHLKPDFLANASRARSQFNLREVSTFRFLPG 492
Qy      481 EYLVSTPEPNKGDPCIRVFSEKADYQAVDEIEANL--EEFDISEDDIDGVRLLFA 539
      493 EYVVSTPEPNKEGDFVLRFPSEKSGATVELDDQIOLNLPDQVLSSEEBIDENFKALFR 552
Qy      540 QLAGEDAISAFELQTLIRRVLAKRQDIDKSGSIFETCKIMVMDLSDSGKGLKEFYI 599
      553 QLAGEMEISVKELRTILNRILISKHDLRTKGSLSBSCSMVNLMDGNGKGLVEFNI 612
Qy      600 LMTKIOYOKIYREIVDRSGTWNYSYEMRKALBEAGFKMPCQHQIVARFADQILIF 659
      613 LMRIRNYLSIFKFPDLDSGMSAYEMRMALISAGFKLNKLYELITRYSERDLAVDF 672
Qy      660 DNFVCLVRLLETFLFKI FKOLDPENTGTIELDLISWL 695
      673 DNFVCLVRLLETFLFRFKTLDTDLDGAVTFDLFKWL 708
Db

```

RESULT 7

```

US-10-116-519-6
; Sequence 6, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-519-6

```

Query Match 66.2%; Score 2458.5; DB 15; Length 714;
 Best Local Similarity 63.1%; Pred. No. 6.3e-221;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

```

Qy      3 GINAKLAKDEALGEGLSHERAITYLNQDYEAALNCELEGLTFOQDSFPAISALGFE 62
      13 GVSADVOKORARELIGRHNENAIKYLGOQDYEBOLRVRCLOSGTLFRDEAFPPVPSISLGYKD 72
Db

```

```

Qy      63 LGFYSSKTRGMKWRKPTTEICADPOFTIGATRTDICGALGDCMLAAIASLTINERILA 122
      73 LGFNSSKTYGIGKKRPTBELSNPQFIVDAGTRTIDICGALGDCMLAAIASLTINTLILH 132
Db      123 RVPPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTQDGLLFFVHSAEGSFMSALLE 182
      133 RVPFHGSGFONGAGIHFHOLMOGEWVDVVDLLPIKDGKLVFHSAGNEFWMSALLE 192
Qy      183 KAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIIOKALOKSILGCSI 242
      193 KAVAKONGSYEALSGSTSGFEDFTGIVTEWELKRAPSDLYOIILKALERSILGCSI 252
Qy      243 DITSADSEAITFOKLVKHAIVTGAEEVNSGLOKLRIRINPGEVMTGRANDNCP 302
      253 DISSVADMEAITFKLVKHAIVTGAQVNVGQVSLIRMRNPMGEVMTGAMSDSS 312
Db      303 SMTIDPEERERLTRRHEDGEFWMSPDLRHYSRLICNLTPDTLSDTYKKKLTMD 362
      313 EMNVVDYERDQLRVKMEDEFPWMSFRDPFRBEFTRLICNLTPDALKSRTIRKNTTLYE 372
Qy      363 GNMRRGSTAGGCNYPNTFMMNPQVYLKLEEDDEE--DGESGCTFVLGLIOKRRROR 420
      373 GTMRSGSTAGGCNYPATFVWNPQFKRLDETDPDYGDRSGCSFVLALMQHRRER 432
Qy      421 KMGEDMHTTGFYIEVPEELSGQTNHLSKNFPLTNRAERSDTFNLREVLNRFKLP 480
      433 RFRDMETTGFAYEVPPELVGQPAVHLKPDFLANASRARSQFNLREVSTFRFLPG 492
Qy      481 EYLVSTPEPNKGDPCIRVFSEKADYQAVDEIEANL--EEFDISEDDIDGVRLLFA 539
      493 EYVVSTPEPNKEGDFVLRFPSEKSGATVELDDQIOLNLPDQVLSSEEBIDENFKALFR 552
Qy      540 QLAGEDAISAFELQTLIRRVLAKRQDIDKSGSIFETCKIMVMDLSDSGKGLKEFYI 599
      553 QLAGEMEISVKELRTILNRILISKHDLRTKGSLSBSCSMVNLMDGNGKGLVEFNI 612
Qy      600 LMTKIOYOKIYREIVDRSGTWNYSYEMRKALBEAGFKMPCQHQIVARFADQILIF 659
      613 LMRIRNYLSIFKFPDLDSGMSAYEMRMALISAGFKLNKLYELITRYSERDLAVDF 672
Qy      660 DNFVCLVRLLETFLFKI FKOLDPENTGTIELDLISWL 695
      673 DNFVCLVRLLETFLFRFKTLDTDLDGAVTFDLFKWL 708
Db

```

RESULT 8

```

US-10-116-519-8
; Sequence 8, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-116-519-8

```

Query Match	65.6%;	Score 2435;	DB 15;	Length 713;
Best Local Similarity	62.8%;	Pred. No. 1e-218;		
Matches 437;	Conservative 126;	Mismatches 129;	Indels 4;	Gaps 3;

Qy	3	GIATLADREAAEELGSHERAIKTYLNDYELARECEAGLTPODSPPALPSALGFEE	62
Db	13	GVSNOVOKRDELELGRHENAIKTLGODYELRLARCLQSGVLFODEAPFPVSHSLGRE	72
Qy	63	LGPYSSKTRGMKRPTEICADPOPIGATRTDICOGLGDCWLLAIASITLNEETIA	122
Db	73	LGPSSKTYGIKMKRPTEILMSNPQIVGATRTDICOGLGDCWLLAIASITLNEETILH	132
Qy	123	RVPLNPSOFOENYAGIFPHFOFYQGEWVVDVDDLPTKDGELLVHASGESEFMSALLE	182
Db	133	RVPLPGSFOGYSAGIFPHFOLMQSEBWDVVDLPLTKDGKLVVYHSAQNEFMSALLE	192
Qy	183	KAYAKINGCYEALSGATTEGEDEFTGGIAEWYELKKEPPNLFKIIORALOKGSLGCSI	242
Db	193	KAYAKVNSYEALSGGCTSEAFEDPTGVTWEYDLOKAPSLDIYIILKALERGSLLGCSI	252
Qy	243	DITSAADSEATTPQULVNGHAYSVYGAEBEVSNGSLQKILIRNPMGEVEMGRWNNDCP	302
Db	253	NISDRIDLEATTFKULVRGHASVYGAOVYTGQGRVNLIRNRPMGEVEMGRWSDSSY	312
Qy	303	SWNTIDPEPERELTRRHEDGEFMSFSDFLRHYSRLKICNLTPTDITSDTYKKMLTKTMD	362
Db	313	EMNKVDYPEREGDLARKMEDGEFMSFRPIRFTLTLEICNLTPLDKSTLTANNMTTFYE	372
Qy	363	GNMRGSGTAGCGRNTPTFMNPOYLKLEED--EDEEDSGGCTFLVGLIOKERRROR	420
Db	373	GTWRGSGTAGCGRNPATFWNPQKIRLEEDVDDADDVNRSGGSFLTALMOKERRER	432
Qy	421	KMGEDMHTIGFIVYPEELSGOTNHLISKNFELNRAERSDTINLREVNARFKLPPG	480
Db	433	RFGEDMETIGFAVYVPREIAGQP-VHLKRPDFLNASTAGSEHNTINLREVSNRIRLPPG	491
Qy	481	EYILVPSFTPEPNKODFCIRVFSEKKADYQAVVDEIEANL-EEFDISEDDIDGVRLLEPA	539
Db	492	EYIVVPSFTPEPNKEDFLIRFSEKKAGQELDDQIOANLPBEKYLSEEDIDNKEFLIFS	551
Qy	540	OLAGDAISAELQTLIRYLAKRODILKSDPSFLETCKIMDMILDSGSGGLGKEFTI	599
Db	552	KLAGDMEISYKELQTLINRIILSKIKDLPTNFSLESCKSVNMLDNDONGLGLIVEFTI	611
Qy	600	LMTKIOKOKIYRELDVDRSGTMSYEMRKALKEBAGFKMPCOLHOVIVARFADDDLIIDE	659
Db	612	LMNRIRYNLTJIFRKFDLDKSGSMSAYEMRMATEAAGFKLKKKLTHELIIIRYSEBPLAVDF	671
Qy	660	DNFPRCLVRLLETLFKIQDPEANTGTIELDLSML	695
Db	672	DNFVCCIVRLLETMFREFKLLDIDDGVYTFDLFKML	707

RESULT 9
US-09-768-877-27
Sequence 27, Application US/09768877
Patent No. US20020150896A1
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELLI, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/768, 877
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 09/422, 869
PRIOR FILING DATE: 1999-10-21

```

: NUMBER OF SEQ. ID NOS: 30
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 27
: LENGTH: 703
: TYPE: PRT
: ORGANISM: RAT
: US-09-768-877-27

Query Match          63.5%; Score 2356.5; DB 10; Length 703;
Best Local Similarity 60.9%; Prod. No. 2.2e-211;
Matches 427; Conservative 127; Mismatches 140; Indels 7; Gaps 3;

```

Oy	1	MAGIAALADOREABEALISGHERIAIKXNDYELANECBEAGLTPODSEFPALPSALGF	60
Db	1	MAALAAVYSXORAVABEALISGNONAVKYLQODFETLRQCLANSVGLFDPDEPPACPSALGY	60
Oy	61	KELGPVSKTRGMWRMKPTEICADPOFIIGATRTDICOALGDWCMLAAIASITLNEBI	120
Db	61	KDLGPGSPDIOGIVMKRPTFLCPNPQGITVGGATRTTDIROGGLGBWCMLAAIASITLNEKL	120
Oy	121	LARVAVLNQSFQENYAGIFHFQEWQGEWEVYVDDRLLPTKDGELLFVHSABESGEFMSAL	180
Db	121	LYRVLPRDQSFQOXDYGAFIHFQFQWQGEWEVYIDDLPTKNGQLFLHSEEGNEFMSAL	180
Oy	181	LEKAYALINGCYELASGATTBEGEDPTGSIAMELYELKCRPNLFIIOALQKQSLIGC	240
Db	181	LEKAYALINGCYELAVGSGTTEGEPDPTGSISEFYDLKCRPNLYIIIOALRGSLLGC	240
Oy	241	SIDTSAADSEALTPQCLVKGHAYSVTGAEEVENSGLQTLIRINPMGEWMTGRANDN	300
Db	241	SIDVSTAAEBATTRQCLVKGHAYSVTGEBVNFHGRPKLIRINPMGEVENSGLMSDN	300
Oy	301	CPSWNTIDPEBERLTERHDEGEFMSFSDFLRHYSRLIECNLTPTDTLTSDYKKWKLTK	360
Db	301	APENNYIDPRKEELDKKADGEFMSFSDFLQYSLEICNLSPDLSSEBIHKNLVL	360
Oy	361	MDGNMRGSGTPOGCRNYPNTFMNPOYLKLEEBEDEBEOG--ESGCTFLVGLIQKRRR	418
Db	361	FNCRWGTSGTAGGCLNPGYVWNTNPOKIHLDVEDEOEBEGTSEPCCTVLLGLMQKRRR	420
Oy	419	QRKKGEMHNTPGIOIYVPEBELSGQNTIHLSKNFELTNRARESDTINLRVYLNRRKLP	478
Db	421	QKRGGQEMLSIGAVALYQIPLELSBHDAHGRDFELGRQSTGSIYMNLRVSSRVLPL	480
Oy	479	PGEYIIVPSTFEPRKODPCIRVSEKKADYQAUWDEIEANLEB---FDISEDDIDDGVR	535
Db	481	PGQYLVVPSTFEPRKODPCIRVSEKKALALEIGDVSQHPHPRHMDDE--DEHVR	538
Oy	536	RLFAQLAGEDAEISAFELQTLIRRVIAKRODIDKSGFSIBTCKIMVMDLSDSGSGKLGK	595
Db	539	SLFEEPAVGKDSISANQDKRVLNEVLSKRTDMKPDGFNINTCREMISILSDSDGSGLGM	598
Oy	596	EFTYLMTRKIQKQKIYREIDVDSGTMNSIEMRKALBEAGFKMPCQLHOIYVARFADDOL	655
Db	599	EFTYLMTRKIYREIIPQEMHNHVGITIEAHMRJALKAGFTLNNVOQOTIAMRYACKL	658
Oy	656	IIDPDNVRCLVRLLETLFKIKOJDPENGTGIELDLSMLC	696
Db	659	GVDPENGFVACMIRLETLFKLFRLLDKQONGIIVQSLAMEWLC	699

```

RESULT 10
US-10-116-519-11
; Sequence 11, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE
; TITLE OF INVENTION: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253

```



```

? PRIOR FILING DATE: 2001-04-03
? PRIOR APPLICATION NUMBER: US 60/288,768
? PRIOR FILING DATE: 2001-05-04
? PRIOR APPLICATION NUMBER: US 60/296,180
? PRIOR FILING DATE: 2001-06-06
? PRIOR APPLICATION NUMBER: US 60/300,620
? PRIOR FILING DATE: 2001-06-25
? NUMBER OF SEQ ID NOS: 145
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 11
? LENGTH: 702
? TYPE: PR1
? ORGANISM: Homo sapiens
US-10-116-519-11

Query Match      54.2%, Score 2014, DB 15, Length 702;
Best Local Similarity 52.4%, Pred. No. 2,4e-179;
Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;

QY 4 IAAATLADREAAEGSGHERAIKTLNDYELARNCEAGCTLFODPSPAIPSLAGFREL 63
D 1 MVAALINNSRLKAKGVGGHDNAQNPNNGNOSFEBELRAACIRKGBELFEDPLPAPAPSSIGFXDL 60
QY 64 GPYSKTRGKMKKPTLEICADPOPIFGATFTDI CQALGBCWTLAALASTINEETIAR 123
D 61 GPNSKAVONISWQPKDIIINPLTMDISPTDICOGLGBCWTLAALAGSLTTCPKLIYR 120
QY 124 VVPLNOSFOENYAGIFHFQFQWQYEWVWVVDRLPTKDGELFVHSAEGSEFMSALIEK 183
D 121 VVPRGQSPFKKVAIGIFHFQIWFQGWVWVVDRLPTKNDGLVYHSTRESEFMSALIEK 180
QY 184 AYAKINGCYEALSGCATTEGDEFTTGGIAEYBELKRPPLPKIIQKALQKSLGCSID 243
D 181 AYALSGSYEALSGGATWEGLEDEFTGGVAOSFOLQRPQNLRLRLRKAVERSISMGCSIE 240
QY 244 ITSADDEALTFQGLVNGHAYSVTGAEBVESNGSLQKILIRHNPGEVETWGRWNNDOPS 303
D 241 VTSSELESMTDKMLVGHAYSVTGLDVAHYRGMKETLIRRNPNPGRLEWNGAWSDSARE 300
QY 304 WNTIDPEPERELTRRHEDGEFWMGSFDFLRHYSRLIECNLTPTDITSDPYKKMLTKMDG 363
D 301 WEVVASDIOQLAKHTEDGEFWMGYODFLANNFTLIEICNLTPDTLSGDYKSYWHTTFYSG 360
QY 364 NWRBGSTAGGCRNYPNTFWMNPQYLLKLEEBDEDEBDESG---CTPLVGLIQHRROR 420
D 361 SWRGSSAGGCRNHPGFTWTPQPKISLPBDDDEDDAEGNVVVCTCLVALMQNMHHAR 420
QY 421 KMGEDMHTIGGITYEVBEELSQCQNIHLSKPFLLTNRAERSDFTINLREVLNPKLPBG 480
D 421 QQGAQLDTIGVLAVALPKEFONIDVHLKKEFFTKYQDHGFSEIFTMSREVSOLRLRPG 480
QY 481 EYIIVPSTPEENKQDPCIRVSESKKADYQAVND-ETEANLEEDDISRDDIDDOVRRLFA 539
D 481 EYIIPSTPEENHRADELRLVFTKHSSEWLEDNVAEOUEKVESDDMDQPLHFK 540
QY 540 QLAEDBAISAFFELDTLLRVLAKRODIKSDGFSIETCKIAMDILDSGSKGLAKEFYI 599
D 541 IVAEGKEIGVVELQRLNRLWAIKFKSKTKTGFGDLARCMINIMLMDXGSKGLLEPKI 600
QY 600 LMTKIQQKYQKTYREIIVDRSGTWNASYEMRKALAEAGFMPCOLHOVIYARFADQLIIDF 659
D 601 LMKLKKKMDIFRECDQDHSSTLNSYEMRLVIERAGIKLNKKVMQVLAARVADDLIIIDF 660
QY 660 DNPFVCLVRLLETLFKIKQALDRENTGTLEDLISML 695
D 661 DSFISCFURLKTMFTFFLMDPKNTGHCISLEQWL 696

RESULT 11
US-10-116-519-9
? Sequence 9, Application US/10116519
? Publication No. US20030114373A1
? GENERAL INFORMATION:

```

APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
TITLE OF INVENTION: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
FILE REFERENCE: D0124 NP
CURRENT APPLICATION NUMBER: US/40/116, 519
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/281, 253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288, 768
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/296, 180
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/300, 620
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
US-10-116-519-9

Query Match 52.3%; Score 1943.5; DB 15; Length 709;
Best Local Similarity 51.0%; Pred. No. 9.8e-173; Indels 7; Gaps 4;
Matches 354; Conservative 136; Mismatches 197;

7 KLANDREAEGLSGHERAIKYNODYEALRNECEAGTLFODPSFPAISALGFKEPLPY 66
12 RVIRERRBRNGEIVSQPLKEGODFVYLKRCCLAKOKLFEEDRVFPAVOALGSHLS-Q 70
67 SSKTRGMRKRPETRICADPOFIGATRTDICGALGDCWLLAASLTNBEILARVP 126
71 KAKKAKATWKKPKKEICENPGFIIGANRTDICGDLGDCWFLAIACTLIERLFRVIP 130
127 LNSGFQENYVGIHFHOFQWQGEWEVVDVDDLPTKDGSLFVHSASESEFWSALLEKAYA 186
131 HDQSFTEYVAGIHFHOFRRYGDWDVVDLDCLPYNNQLVFTKSNHREFFWSALLEKAYA 190
187 KINCCYEALSGATTEGFEFTGGIAEWYELKKEPPNLFKTIQALQKSGLSGCSIDITS 246
191 KLHGSYELKKGNTTEMEDFTGVTFFELKADAPSDMYKIMRKRIENGSLMGCSIDITV 250
247 AADSEAITFOKLTVKGHAYSVTGAEEVNSGLQKILIRINMGVEYEWGTGRWNNDCPSMNT 306
251 PVQYETRMACGLVGHAYSVTGLEALFKGEKVLVRLRNPMGVENMGWSMDGKMSF 310
307 IDPERERELTRR-HEDGEFWMSSDFLRHYSRLICNLTPTLTISDYTKKAKLTKMGNW 365
311 VDKDEKARLQHVLEDEGFWMSSYDDFVYHFKLEICNLTALAESDRLQTTWVSVNEGRW 370
366 RRGSTAGGCRNYPNTFMMNPQYLLKLEBEDDEDESGSCFTVLGLIQKHRRORKMED 425
371 VRGSGAGGCRNFPPTFMTNPQYRLKLEKEDDPEDESEYCSFLVAAKMDKRRKDKLQAN 430
426 MHTTIGFIYVEPELISGQTNILHLSKNFFLTNRARESDPTINLREVLNFKLPGEXYLIV 485
431 LFTIGFALIVPEPKMHGNKQ-HLQKDFPLVYASARSSTYINMREVSGRFLPPEYIV 489
486 PSTPEPNKDGECIRVFSEKADYQAVDELEAN---LEEFIDSEDDIDGVRRLPAQL 541
490 PSTPEPNKDGECIRVFSEKADYQAVDELEAN---LEEFIDSEDDIDGVRRLPAQL 549
542 AGEBAELISAPFLQITLIRVLAKRODIDSQGISITCKIMVMMLSDSGSKLGLKEFYLLM 601
550 AGDMEICADLKNVLTIVNKKHDLKTQGGTLLSECSRMALMLPDDSGRLNLOEFHLLW 609
602 TKIQYKQIYAEIIVDSSGTNNSYEMRKALEEAEAFKMCQQLHQYIVAFADQDLIDPDN 661
610 KKIYAKMGKIFGHVYDTHSGTINSYEMRNAYVADAFHLSQLYDITIRYADKXHNMDPDS 669
662 FVRCLVRLLETLFKIKQLDPDENTGTIELDLISWL 695
670 FICCFVRLGEGFRAFNAPDKDGDGIKLTNVLLEWL 703

QY 507 ADVADDEEALNE-----FDISEDDIDGVRRLPAQLAGDAEISAFELQTLIRVL 561
 Db 491 AITREDMGNVVIDIDPEEPKPTPPDQTEBEEQFRALFEQVAGEDEMEVTAELVEYVLA 550
 QY 562 AKQDIDSDFSIETCKIMVMDLSDSGKGLKEPFIYLMWKIOXKIXEIVDVSCT 621
 Db 551 OKKQDIFPKLISLISCKNITISLMDTSGKLEPDEFPMKLOMNLPLRFPADSGT 610
 QY 622 MNSYEMKALAEAGFKAPCOLHQYIVARFADQOLIIFDNEFVRCVLETLTFIKIQLDP 681
 Db 611 MSTEALATLKAAPQOLSHLLQIVLRKXADELQDPDDCLNCLVLENNASRYFQALST 670
 QY 682 ENTGTIELDL 691
 Db 671 KNEKEFIHLNI 680

RESULT 14

US-09-768-877-24
 ; Sequence 24, Application US/09768877
 ; Patent No. US20020150896A1
 ; GENERAL INFORMATION:
 ; APPLICANT: POLONSKY, KENNETH S.
 ; APPLICANT: HORIKAWA, YUKIO
 ; APPLICANT: ODA, NAOHISA
 ; APPLICANT: COX, NANCY J.
 ; APPLICANT: SREENAN, SEAMUS
 ; APPLICANT: ZHOU, YUN-PING
 ; APPLICANT: OTANI, KENICHI
 ; APPLICANT: HANIS, CRAIG L.
 ; APPLICANT: BELL, GRAEME I.
 ; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
 ; FILE REFERENCE: ARCD:307
 ; CURRENT APPLICATION NUMBER: US/09/768, 877
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 09/422, 869
 ; PRIOR FILING DATE: 1999-10-21
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 821
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-768-877-24

Query Match 50.2%; Score 1862.5; DB 10; Length 821;
 Best Local Similarity 45.5%; Pred. No. 4.7e-165; Indels 105; Gaps 8;

Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;

QY 32 YEALRNECEAGTLFODPSFPALPSALGFELGPFYSKTRGMKRPTEICADPOFTIG 91
 Db 61 FEOHKKCLEKVLVYVDEPPDETSLFYSGKPIQ-----FWMKRPEICENPRFTIG 115
 QY 92 ATRPTDIOGALGDCWMLAIAIASLTINEELARVPLNOSQENYAGIFHFQWQYGEWE 151
 Db 116 ANRTDIOGELGDCWFLAIAIACLTINOHLLFRVIPHQSFTEYVAGIFHFQWRYGEMVD 175
 QY 152 VVVDRLPTKDGELLFVHSAEGSEFMSALEKAYAKINGCEALSGATTEGEPDFGCI 211
 Db 176 VVIDDCPLTYNNQVLFYKSHRNEFWMSALEKAYAKIAGHSEALKGANTTEAMEDFTGCV 235
 QY 212 AEWEYLKPPPNLFKIIQALOKGSLGCSID-----ITSADSEAITFOK----- 257
 Db 236 AEFPEINDAPSDMKIKKAIERGSILMGCSIDDTNMTYGTSPGLMMGELIAMVNMMD 295
 QY 258 -----LVKGHAYVTGAEBVESNGSLQKILR 283
 Db 296 NSLLQDSDLPDGRSDERPTRTIIPQYETRMACGLVRGHAYSVTGLDEVFPKGEKVLVR 355
 QY 342 IRNMGAEVMTGRANDOCPSMNTIDPERERLTRR-HEDEGFNMSFSDPLRHYSRLEICN 342
 Db 356 LRNMGVENVNGMSWSDKMSFVQKDEKARLQHOVTEDEGEFMSYEDFIYHFKLEICN 415

QY 343 LTPDPTLSDTYKKMKLTKRDGNRRSGTAGCCNNYPNTFMNPOYLIKLEEDDEEDGE 402
 Db 416 LFDALQSDQLQWTWVSVNGRWVRCGACGNFPDTEFWTNPOYRLKLEEDDDPDSE 475
 QY 403 SCGTFVLGLQKRRRRKRGEDMHTGPGIYVPELSQTMNHLKSNFLLNRARERS 462
 Db 476 VLSFVLALMQKRRKDRKLGASLFTIGFALYEPKEMHGNKO-HQKDFPLVNASYARS 534
 QY 463 DTFINLEEVNRRFLPGEIYLVSPPEPKDQDFCIRSEKKADQOAVDEIEANTL-- 520
 Db 535 KTYINRREVQRRLRPPSEVYIVPSTYEPHQBSEFLIRVSEKRNLSSEVENTISVDRPV 594
 QY 521 -----EEFDISBDIDG----- 533
 Db 595 KKKTKPIIFVSDRANSKELGVQDS-EEGKGKTPSPDKQSPQPPQSGSDQSEBQQ 653
 QY 534 VRLFAQAGEDAEISAFELQTLIRRLARQDIDSDEIEFCKIMVMDLSDSGKLG 593
 Db 654 FRNIFKQIAGDDMEICADELKKVLTWNHGXKDKTHGFTLESCRSMIALMDTDSGKLN 713
 QY 594 LKEFYIIMTKIOXKOKIYREIDVDRSQTMNSYEMKALAEAGFKMPCOLHQVIVARFAD 653
 Db 714 LOEFHLMKIKAMQKIFKHYVDQSGTINSYEMKNAVNDAGFHLNQLYDITMRADK 773
 QY 654 QLIIDFQNVCLVRLTEFLFKIKQLDPENTGTIELDLISWL 695
 Db 774 HNMIDFDSFICFVRLEGWFRFAHFDKDGDIKLVLEWL 815

RESULT 15

US-10-116-519-7
 ; Sequence 7, Application US/10116519
 ; Publication No. US2003014373A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
 ; FILE REFERENCE: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
 ; CURRENT APPLICATION NUMBER: US/10/116, 519
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US 60/281,253
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/288,768
 ; PRIOR FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/296,180
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/300,620
 ; PRIOR FILING DATE: 2001-06-25
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 821
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-116-519-7

Query Match 50.2%; Score 1862.5; DB 15; Length 821;
 Best Local Similarity 45.5%; Pred. No. 4.7e-165; Indels 105; Gaps 8;

Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;

QY 32 YEALRNECEAGTLFODPSFPALPSALGFELGPFYSKTRGMKRPTEICADPOFTIG 91
 Db 61 FEOHKKCLEKVLVYVDEPPDETSLFYSGKPIQ-----FWMKRPEICENPRFTIG 115
 QY 92 ATRPTDIOGALGDCWMLAIAIASLTINEELARVPLNOSQENYAGIFHFQWQYGEWE 151
 Db 116 ANRTDIOGELGDCWFLAIAIACLTINOHLLFRVIPHQSFTEYVAGIFHFQWRYGEMVD 175
 QY 152 VVVDRLPTKDGELLFVHSAEGSEFMSALEKAYAKINGCEALSGATTEGEPDFGCI 211
 Db 176 VVIDDCPLTYNNQVLFYKSHRNEFWMSALEKAYAKIAGHSEALKGANTTEAMEDFTGCV 235
 QY 212 AEWEYLKPPPNLFKIIQALOKGSLGCSID-----ITSADSEAITFOK----- 257

```
Db 236 AEFPEIRDPADMYKIMKKAIERGSLMGCSIDGTMVTGTSPSGLNMGELIARVNRMD 295
QY 258 -----LVKGHAYSVTGAEEVESNGSLQKLR 283
Db 296 NSLIQSDLDPRGSDERPTTIIPOVETRNACGLVRGHAYSVTGLDEVPRKGEKVLVR 355
QY 284 IRNPGVEYEWTCRWNDNCPSWNTIDPEERERLTRR-HEDGEFMSFSDFLRHYSRLICN 342
Db 356 LRNPWGQYEWNGSWSDBRKDWSFVDKDEKARLQHQVTEDEGEFMSYEDFIYHFTKLEICN 415
QY 343 LTPDTLSDTYKWKLTMDGNWRGSLAGCCRNVPNTFMNPOYLKLEBEDDEDEGE 402
Db 416 LTAADALQSDKLOTWTVSVNEGRWVRGCSAGCRNFPDTEFTWNPQYRLKLEEDDDPDDSE 475
QY 403 SGCFELVGLIOXHRBROKMGEDMHTIGFGIYEVPEELSGQTNTHLSKNPFLTRARERS 462
Db 476 VICSFLVALMOKNRKDKRLGASLFTIGFALYEVPEKMHGKQ-HLQKDFLYNASKARS 534
QY 463 DTFINLREVLNRFKLPRGEYILVSTPEPNKDGFCIRVSEKKADYQAVDEIEANL-- 520
Db 535 KTYINMRVSGRFRLPSEYIVPSTYEPHQGEFILRVFSEKNTLSEVENTISVDRPV 594
QY 521 -----EEFDISEDDIDG----- 533
Db 595 KKKKTKPIIFVSDRANSNKEIGVDQES-EEGKGKTSPPDKQKQSPQPPGSSDQSEEQQ 653
QY 534 VRRLFAOLAGDEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVDM.LSDSGSKLG 593
Db 654 FRNIFKQIAGDDMEICADELKKVLTWVVKHDLKTHGFTLSCRSMTALMDTDGSKLN 713
QY 594 LKEFYILMTKLOXYOKIYREIDVDRSGTMSYEMRKALEBAGFKMPCQLHQVIVARFAD 653
Db 714 LQEFHLLMNKIKAMQKIFKHVDTOGSGTINSYEMRNNAVNDAGFHLNNQLYDIITRVYDX 773
QY 654 QLIIDFNFRCLVRLFTLFIKIFKQIDPENTGTLELDLSWL 695
Db 774 HMNIDFDSFICCFVLEGMFRFAFADKDGDIKILNVLWL 815
```

Search completed: July 24, 2003, 13:14:56
Job time : 50.4084 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:48:39 / Search time 30.9606 Seconds
(without alignments)
956.622 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MAGIAAKAKDRAEAGLS.....PENTGTIELDLISMLCFSVL 700

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	US-08-726-525-7	Sequence 7, Appli
2	3713	100.0	700	US-08-487-942-7	Sequence 7, Appli
3	3713	100.0	700	US-08-726-036A-7	Sequence 7, Appli
4	3713	100.0	700	US-09-422-869-23	Sequence 23, Appli
5	3713	100.0	700	US-09-083-516-7	Sequence 7, Appli
6	3491	94.0	700	US-09-308-345A-46	Sequence 46, Appli
7	2458.5	66.2	714	US-09-422-869-22	Sequence 22, Appli
8	2447.5	65.9	714	US-09-308-345A-47	Sequence 47, Appli
9	2370.5	63.8	703	US-08-835-099A-1	Sequence 1, Appli
10	2370.5	63.8	703	US-09-157-349-1	Sequence 1, Appli
11	2370.5	63.8	712	US-08-835-099A-2	Sequence 2, Appli
12	2370.5	63.8	712	US-09-157-349-2	Sequence 2, Appli
13	2356.5	63.5	703	US-09-422-869-27	Sequence 27, Appli
14	2308.5	62.2	703	US-09-308-345A-49	Sequence 49, Appli
15	2020	54.4	739	US-09-653-839-2	Sequence 2, Appli
16	2014	54.2	702	US-09-653-839-6	Sequence 6, Appli
17	1967	53.0	723	US-09-653-839-4	Sequence 4, Appli
18	1961	52.8	686	US-09-653-839-8	Sequence 8, Appli
19	1925.5	51.9	757	US-09-622-869-1	Sequence 1, Appli
20	1882	50.7	690	US-09-422-869-28	Sequence 28, Appli
21	1869.5	50.4	821	US-09-622-880B-15	Sequence 15, Appli
22	1862.5	50.2	821	US-09-422-869-24	Sequence 24, Appli
23	1839.5	49.5	821	US-09-308-345A-48	Sequence 48, Appli
24	1594.5	34.9	828	US-09-308-345A-19	Sequence 19, Appli
25	1287	34.7	251	US-08-726-525-6	Sequence 6, Appli
26	1287	34.7	251	US-08-487-942-6	Sequence 6, Appli
27	1287	34.7	251	US-08-726-036A-6	Sequence 6, Appli

28	1287	34.7	251	4	US-09-083-516-6	Sequence 6, Appli
29	965.5	26.0	640	4	US-09-308-345A-2	Sequence 2, Appli
30	961.5	25.9	513	4	US-09-308-345A-4	Sequence 4, Appli
31	917	24.7	640	4	US-09-308-345A-7	Sequence 7, Appli
32	901.5	24.3	639	4	US-09-422-869-25	Sequence 25, Appli
33	789	21.2	648	4	US-09-308-345A-20	Sequence 20, Appli
34	780	21.0	641	3	US-09-422-869-18	Sequence 18, Appli
35	621.5	16.7	666	3	US-09-422-869-16	Sequence 16, Appli
36	607	16.3	544	3	US-09-422-869-4	Sequence 4, Appli
37	604	16.3	672	3	US-09-422-869-2	Sequence 2, Appli
38	599	16.1	517	3	US-09-422-869-8	Sequence 8, Appli
39	575	15.5	517	3	US-09-422-869-6	Sequence 6, Appli
40	531	14.3	444	3	US-09-422-869-10	Sequence 10, Appli
41	458	12.3	266	3	US-09-032-523-7	Sequence 7, Appli
42	450	12.1	268	3	US-08-835-099A-9	Sequence 9, Appli
43	450	12.1	268	3	US-09-157-349-9	Sequence 9, Appli
44	423	11.4	248	3	US-09-032-523-1	Sequence 1, Appli
45	304.5	8.2	854	2	US-08-928-692-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-08-726-525-7
Sequence 7, Application US/08726525
Patent No. 5789181
GENERAL INFORMATION:
APPLICANT: Lin, Lin-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,525
FILING DATE: 07-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,942
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-726-525-7
Query Match 100.0%, Score 3713, DB 1, Length 700,
Best Local Similarity 100.0%, Pred. No. 0,
Matches 700, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

```

QY 1 MAGIAAKLAKDREAABGLSHERAIKYINDYALRNECLAGTLFQDPSPFPAIPSLG 60
D 1 MAGIAAKLAKDREAABGLSHERAIKYINDYALRNECLAGTLFQDPSPFPAIPSLG 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALDCCWLLAIAASLTINEE 120
D 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALDCCWLLAIAASLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFMOYGEWVAVDDRLPTDGGELLFVHSAESSEFMSAL 180
D 121 LARVPLNOSFOENYAGIFHFQFMOYGEWVAVDDRLPTDGGELLFVHSAESSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEDFTGIAEWELKRPPLFKIIQKALQKSLG 240
D 181 LEKAYAKINGCYEALSGATTEGFEDFTGIAEWELKRPPLFKIIQKALQKSLG 240
QY 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVNSGLOKLIIRNPMGEVETGRWMDN 300
D 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVNSGLOKLIIRNPMGEVETGRWMDN 300
QY 301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDLTSDTYKKMULTK 360
D 301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDLTSDTYKKMULTK 360
QY 361 MDGNMRRGSTAGCCRNYPNTFMNPOVLIKLEBEDDEDESGCTFLVGLIQKHRROR 420
D 361 MDGNMRRGSTAGCCRNYPNTFMNPOVLIKLEBEDDEDESGCTFLVGLIQKHRROR 420
QY 421 KMGEDMHTIGFIYEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
D 421 KMGEDMHTIGFIYEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
QY 481 EYILVPSFEPNKGDFCIRVFSEKKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
D 481 EYILVPSFEPNKGDFCIRVFSEKKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMVMDLSDSGKLKKEFYIL 600
D 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMVMDLSDSGKLKKEFYIL 600
QY 601 WTKIQKOKIYREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVYVARPADOLIIDFD 660
D 601 WTKIQKOKIYREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVYVARPADOLIIDFD 660
QY 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCFSVL 700
D 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCFSVL 700

```

```

RESULT 2
US-08-487-942-7
Sequence 7, Application US/08487942
Patent No. 5817476
GENERAL INFORMATION:
APPLICANT: Lin, Lih-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,942
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-487-942-7

```

```

Query Match 100.0%; Score 3713; DB 2; Length 700;
Best Local Similarity 100.0%; Pred. No.0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGIAAKLAKDREAABGLSHERAIKYINDYALRNECLAGTLFQDPSPFPAIPSLG 60
D 1 MAGIAAKLAKDREAABGLSHERAIKYINDYALRNECLAGTLFQDPSPFPAIPSLG 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALDCCWLLAIAASLTINEE 120
D 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALDCCWLLAIAASLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFMOYGEWVAVDDRLPTDGGELLFVHSAESSEFMSAL 180
D 121 LARVPLNOSFOENYAGIFHFQFMOYGEWVAVDDRLPTDGGELLFVHSAESSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEDFTGIAEWELKRPPLFKIIQKALQKSLG 240
D 181 LEKAYAKINGCYEALSGATTEGFEDFTGIAEWELKRPPLFKIIQKALQKSLG 240
QY 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVNSGLOKLIIRNPMGEVETGRWMDN 300
D 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVNSGLOKLIIRNPMGEVETGRWMDN 300
QY 301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDLTSDTYKKMULTK 360
D 301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDLTSDTYKKMULTK 360
QY 361 MDGNMRRGSTAGCCRNYPNTFMNPOVLIKLEBEDDEDESGCTFLVGLIQKHRROR 420
D 361 MDGNMRRGSTAGCCRNYPNTFMNPOVLIKLEBEDDEDESGCTFLVGLIQKHRROR 420
QY 421 KMGEDMHTIGFIYEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
D 421 KMGEDMHTIGFIYEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
QY 481 EYILVPSFEPNKGDFCIRVFSEKKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
D 481 EYILVPSFEPNKGDFCIRVFSEKKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMVMDLSDSGKLKKEFYIL 600
D 541 LAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMVMDLSDSGKLKKEFYIL 600
QY 601 WTKIQKOKIYREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVYVARPADOLIIDFD 660
D 601 WTKIQKOKIYREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVYVARPADOLIIDFD 660
QY 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCFSVL 700
D 661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCFSVL 700

```

RESULT 3
US-08-726-036A-7
Sequence 7, Application US/08726036A
Patent No. 5981482
GENERAL INFORMATION:
APPLICANT: Lin, Lin-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
TITLE OF INVENTION: BINDING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
City: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,036A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-726-036A-7

Query Match 100.0%; Score 3713; DB 2; length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60
1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60
DB 1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60

QY 61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
DB 61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120

QY 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180
121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180

QY 181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240
181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240
DB 181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240

QY 241 SIDTSAADSEAITFOQLVVGHAIVTGAEBVESNGSLQYLIRINPWGEVETGRWNN 300
241 SIDTSAADSEAITFOQLVVGHAIVTGAEBVESNGSLQYLIRINPWGEVETGRWNN 300
DB 241 SIDTSAADSEAITFOQLVVGHAIVTGAEBVESNGSLQYLIRINPWGEVETGRWNN 300

QY 301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHSLEICNLTPTDILTSDTYKKWLT 360
301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHSLEICNLTPTDILTSDTYKKWLT 360
DB 301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHSLEICNLTPTDILTSDTYKKWLT 360

QY 361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEBDESGCTFLVGLIQHRROR 420
361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEBDESGCTFLVGLIQHRROR 420
DB 361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEBDESGCTFLVGLIQHRROR 420

QY 421 KMGEDMTTGFGIYVDEEISGOTNHLKSNFELTPARARSTPTNLREVLNRFKLP 480
421 KMGEDMTTGFGIYVDEEISGOTNHLKSNFELTPARARSTPTNLREVLNRFKLP 480
DB 421 KMGEDMTTGFGIYVDEEISGOTNHLKSNFELTPARARSTPTNLREVLNRFKLP 480

QY 481 EYILVSTFEPNKDGFICRVFSEKADYQAVDDEIANLEEDISEDIDGVRRLFAQ 540
481 EYILVSTFEPNKDGFICRVFSEKADYQAVDDEIANLEEDISEDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKDGFICRVFSEKADYQAVDDEIANLEEDISEDIDGVRRLFAQ 540

QY 541 LAGEDAISAFAELQTLIRVLAARODIKSDFSIEFTCKIWMDLSDSGSLGLKEFY 600
541 LAGEDAISAFAELQTLIRVLAARODIKSDFSIEFTCKIWMDLSDSGSLGLKEFY 600
DB 541 LAGEDAISAFAELQTLIRVLAARODIKSDFSIEFTCKIWMDLSDSGSLGLKEFY 600

QY 601 WTKIQKQKTYREIIVDRSGTMSYEMRKLEAGFMPQLHVIARPADQOLIIFD 660
601 WTKIQKQKTYREIIVDRSGTMSYEMRKLEAGFMPQLHVIARPADQOLIIFD 660
DB 601 WTKIQKQKTYREIIVDRSGTMSYEMRKLEAGFMPQLHVIARPADQOLIIFD 660

QY 661 NFVCLVRLLETLFKIFKOLDPENTGTIELDLISMLCFSYL 700
661 NFVCLVRLLETLFKIFKOLDPENTGTIELDLISMLCFSYL 700
DB 661 NFVCLVRLLETLFKIFKOLDPENTGTIELDLISMLCFSYL 700

RESULT 4
US-09-422-869-23
Sequence 23, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEANUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG I.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD-307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-422-869-23

Query Match 100.0%; Score 3713; DB 3; length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60
1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60
DB 1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECELEAGTLFODPSFPAIPALGF 60

QY 61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
DB 61 KEIGPYSKTRGKRWKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120

QY 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180
121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFWSAL 180

QY 181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240
181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240
DB 181 LERAVAKINGCYEALSGATTEGFEDFTGGIAEWYELKKPPNLFKIQALOKGSLG 240


```

QY 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVNSGSLQKILIRINPMGEVMTGRWNND 300
DB 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVNSGSLQKILIRINPMGEVMTGRWNND 300
QY 301 CPSPWNTIDPEERERLTRRHEDGEFWMGSDPLRIHYSRLICNTLPDITLSPTYKKMLTK 360
DB 301 CPSPWNTIDPEERERLTRRHEDGEFWMGSDPLRIHYSRLICNTLPDITLSPTYKKMLTK 360
QY 361 MDGWRRGSTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
DB 361 MDGWRRGSTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
QY 421 KMGEDMTHTGFGIYEVEEELSGQTNHLKSNFPLTNRARSDFINLRVLANFKLPFG 480
DB 421 KMGEDMTHTGFGIYEVEEELSGQTNHLKSNFPLTNRARSDFINLRVLANFKLPFG 480
QY 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVNDEIEANLEPDISDDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVNDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
QY 601 WTKIQYQKIYREIDVRSSTGTMNSYEMRKALEAGFMPCOLHOVIYARFADQLIIDFD 660
DB 601 WTKIQYQKIYREIDVRSSTGTMNSYEMRKALEAGFMPCOLHOVIYARFADQLIIDFD 660
QY 661 NFVRCVLRLETLFKIFKQDPENTGTTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQDPENTGTTIELDLISWLCFSVL 700

```

RESULT 5

US-09-083-516-7

Sequence 7, Application US/09083516

Patent No. 6300086

GENERAL INFORMATION:

APPLICANT: Lin, Lih-Ling

APPLICANT: Graham, James

TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR

TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: USA

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/083,516

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,942

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

REFERENCE/DOCKET NUMBER: G15258

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 700 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
US-09-083-516-7

```

```

Query Match          100.0%; Score 3713; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGIAKLANDRAPAEGLSHERAIKYLANDYALNNECLAEAGTLFQDSFPALIPALGF 60
DB 1 MAGIAKLANDRAPAEGLSHERAIKYLANDYALNNECLAEAGTLFQDSFPALIPALGF 60
QY 61 KELGYPSSKTRGMRWRPTEICADPOITIGGATRTICQALGDCWMLAIAITLNEE1 120
DB 61 KELGYPSSKTRGMRWRPTEICADPOITIGGATRTICQALGDCWMLAIAITLNEE1 120
QY 121 LARVPLANGSFQENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSAGSEFMSAL 180
DB 121 LARVPLANGSFQENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSAGSEFMSAL 180
QY 181 LERAYAKINGCYEALSGGATTEGPEFTGIAEWEYELKPPNLFKIQALOKGSLGC 240
DB 181 LERAYAKINGCYEALSGGATTEGPEFTGIAEWEYELKPPNLFKIQALOKGSLGC 240
QY 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVNSGSLQKILIRINPMGEVMTGRWNND 300
DB 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVNSGSLQKILIRINPMGEVMTGRWNND 300
QY 301 CPSPWNTIDPEERERLTRRHEDGEFWMGSDPLRIHYSRLICNTLPDITLSPTYKKMLTK 360
DB 301 CPSPWNTIDPEERERLTRRHEDGEFWMGSDPLRIHYSRLICNTLPDITLSPTYKKMLTK 360
QY 361 MDGWRRGSTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
DB 361 MDGWRRGSTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
QY 421 KMGEDMTHTGFGIYEVEEELSGQTNHLKSNFPLTNRARSDFINLRVLANFKLPFG 480
DB 421 KMGEDMTHTGFGIYEVEEELSGQTNHLKSNFPLTNRARSDFINLRVLANFKLPFG 480
QY 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVNDEIEANLEPDISDDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVNDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
QY 601 WTKIQYQKIYREIDVRSSTGTMNSYEMRKALEAGFMPCOLHOVIYARFADQLIIDFD 660
DB 601 WTKIQYQKIYREIDVRSSTGTMNSYEMRKALEAGFMPCOLHOVIYARFADQLIIDFD 660
QY 661 NFVRCVLRLETLFKIFKQDPENTGTTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQDPENTGTTIELDLISWLCFSVL 700

```

RESULT 6

US-09-308-345A-46

Sequence 46, Application US/09308345A

Patent No. 6569665

GENERAL INFORMATION:

APPLICANT: BOEHM, Thomas;

APPLICANT: DEAR, Neil T.

TITLE OF INVENTION: No. 6569665el calpains, their preparation and use

FILE REFERENCE: 0050/47576

CURRENT APPLICATION NUMBER: US/09/308,345A

FILING DATE: 1999-05-19

NUMBER OF SEQ ID NOS: 49

SOFTWARE: Wordperfect v. 6.1

SEQ ID NO 46
 LENGTH: 700
 TYPE: PRT
 ORGANISM: mouse
 US-09-308-345A-46

Query Match 94.0%; Score 3491; DB 4; Length 700;
 Best Local Similarity 92.9%; Pred. No. 0;
 Matches 650; Conservative 25; Mismatches 25; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDREAAGLSHERAIKYLNDYALRNECEAGTLFODPSFPAIPALGF 60
 DB 1 MAGIAKLAKDREAAGLSHERAIKYLNDYALRNECEAGTLFODPSFPAIPALGF 60
 QY 61 KELGYSKTRGMRKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASLTINBEI 120
 DB 61 KELGYSKTRGMRKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASLTINBEI 120
 QY 121 LAHVPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 180
 DB 121 LAHVPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 180
 QY 181 LEKAVKINGCYALSGATTEGPEDEFTGGIAEMWELKKRPNNLFKIIQALOKSLGCG 240
 DB 181 LEKAVKINGCYALSGATTEGPEDEFTGGIAEMWELKKRPNNLFKIIQALOKSLGCG 240
 QY 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPMGEVMTGRNND 300
 DB 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPMGEVMTGRNND 300
 QY 301 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 360
 DB 301 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 360
 QY 361 MDNMRGSAAGCGRVNTFMNPOYLKLEEDDEDESGCTVLVGLIOKRROR 420
 DB 361 MDNMRGSAAGCGRVNTFMNPOYLKLEEDDEDESGCTVLVGLIOKRROR 420
 QY 421 KMGEDHTTIGFYEVPEELSGQTNILSKNFLLTNARERSPTFINLRVLRFKLP 480
 DB 421 KMGEDHTTIGFYEVPEELSGQTNILSKNFLLTNARERSPTFINLRVLRFKLP 480
 QY 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLBEFDDISDDIDGVRRLFAQ 540
 DB 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLBEFDDISDDIDGVRRLFAQ 540
 QY 541 LAEDEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKINVDMLDSGSGKLKERYIL 600
 DB 541 LAEDEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKINVDMLDSGSGKLKERYIL 600
 QY 601 WTKIOKQKIYREIDVDSGTNMSYEMRKALBEAGFMPQCLHQVIVARFADDLIIDF 660
 DB 601 WTKIOKQKIYREIDVDSGTNMSYEMRKALBEAGFMPQCLHQVIVARFADDLIIDF 660
 QY 661 NFVRLVRLTETLFKIFKQIDPENTGTIELDLISMLCSVL 700
 DB 661 NFVRLVRLTETLFKIFKQIDPENTGTIELDLISMLCSVL 700

RESULT 7
 US-09-422-869-22
 Sequence 22, Application US/09422869
 Patent No. 6235481
 GENERAL INFORMATION:
 APPLICANT: POLONSKY, KENNETH S.
 APPLICANT: HORIKAWA, YUKIO
 APPLICANT: ODA, NAOHISA
 APPLICANT: COX, NANCY J.
 APPLICANT: SREENAN, SEAMUS
 APPLICANT: ZHOU, YUN-PING
 APPLICANT: OTANI, KENICHI
 APPLICANT: HANIS, CRAIG L.
 APPLICANT: BELL, GRAEME I.

TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
 FILE REFERENCE: ARCD:307
 CURRENT APPLICATION NUMBER: US/09/422,869
 CURRENT FILING DATE: 1999-10-21
 EARLIER APPLICATION NUMBER: 60/134,175
 EARLIER FILING DATE: 1999-05-13
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 22
 LENGTH: 714
 TYPE: PRT
 ORGANISM: Human
 US-09-422-869-22

Query Match 66.2%; Score 2458.5; DB 3; Length 714;
 Best Local Similarity 63.1%; Pred. No. 1,5e-236;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

QY 3 GIAKLAKDREAAGLSHERAIKYLNDYALRNECEAGTLFODPSFPAIPALGF 62
 DB 13 GVAQVOKORARELIGRHNALIKYLGQVEQLRVRCLOSGTLFRDEAFPPQSIGYKD 72
 QY 63 LGYSKTRGMRKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASLTINBEI 122
 DB 73 LGYSKTRGMRKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASLTINBEI 122
 QY 123 RVPPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 182
 DB 123 RVPPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 182
 QY 133 RVPPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 192
 DB 133 RVPPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 192
 QY 183 KAVAKINGCYALSGATTEGPEDEFTGGIAEMWELKKRPNNLFKIIQALOKSLGCG 242
 DB 183 KAVAKINGCYALSGATTEGPEDEFTGGIAEMWELKKRPNNLFKIIQALOKSLGCG 242
 QY 243 DITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPMGEVMTGRNND 302
 DB 243 DITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPMGEVMTGRNND 302
 QY 303 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 362
 DB 303 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 362
 QY 363 GMRGSAAGCGRVNTFMNPOYLKLEEDDEDESGCTVLVGLIOKRROR 420
 DB 363 GMRGSAAGCGRVNTFMNPOYLKLEEDDEDESGCTVLVGLIOKRROR 420
 QY 421 KMGEDHTTIGFYEVPEELSGQTNILSKNFLLTNARERSPTFINLRVLRFKLP 480
 DB 421 KMGEDHTTIGFYEVPEELSGQTNILSKNFLLTNARERSPTFINLRVLRFKLP 480
 QY 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLBEFDDISDDIDGVRRLFAQ 540
 DB 481 EYLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLBEFDDISDDIDGVRRLFAQ 540
 QY 540 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKINVDMLDSGSGKLKERYIL 599
 DB 540 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKINVDMLDSGSGKLKERYIL 599
 QY 600 WTKIOKQKIYREIDVDSGTNMSYEMRKALBEAGFMPQCLHQVIVARFADDLIIDF 659
 DB 600 WTKIOKQKIYREIDVDSGTNMSYEMRKALBEAGFMPQCLHQVIVARFADDLIIDF 659
 QY 660 NFVRLVRLTETLFKIFKQIDPENTGTIELDLISMLCSVL 708
 DB 660 NFVRLVRLTETLFKIFKQIDPENTGTIELDLISMLCSVL 708

RESULT 8
 US-09-308-345A-47
 Sequence 47, Application US/09308345A
 Patent No. 656965
 GENERAL INFORMATION:
 APPLICANT: BOEHM, THOMAS;

APPLICANT: DEAR, Neil T.
TITLE OF INVENTION: No. 65696561 calpains, their preparation and use
FILE REFERENCE: 0050/47576
CURRENT APPLICATION NUMBER: US/09/308,345A
CURRENT FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: WordPerfect v. 6.1
SEQ ID NO 47
LENGTH: 714
TYPE: PRT
ORGANISM: human
US-09-308-345A-47

Query Match 65.9%; Score 2447.5; DB 4; Length 714;
Best Local Similarity 62.9%; Pred. No. 1.9e-235;
Matches 436; Conservative 124; Mismatches 131; Indels 3; Gaps 2;

3 GIAAKLADREAAAGLSHERAIKYLNDYALNCEIAETLRQDSFPALPSALGFE 62
13 GVSAGVQKORARELGLGRHENAIXYLGDYEQLRKCLQSGTLPRDEAFPPVPSLGKD 72
63 LGPSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAASLTINEEL 122
73 LGPSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAASLTINEEL 132
123 RVVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFWMSAL 182
133 RVVPHGSGFONGYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFWMSAL 192
183 KAVKINGCYALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILGCSI 242
193 KAVKINGCYALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILGCSI 252
243 DITSAASEATLTKLVKGAHAYVTGAEEVNSGLQKLIIRPMWGEVMTGRMNCP 302
253 DISVLDEALTEFKLVKGAHAYVTGAEEVNSGLQKLIIRPMWGEVMTGRMNCP 312
303 SWNTIDEEERELTRRHEDGEFWMSPDFLRHYSRLICNLTPDLTSDYKKWKLTKMD 362
313 EMNVVDYERDQKLVKMGDEGFWMSPDFLRHYSRLICNLTPDLTSDYKKWKLTKMD 372
363 GNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEB--DESGCTFLVGLIQHRRR 420
373 GPMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEB--DESGCTFLVGLIQHRRR 432
421 KMGEDMHTIGFGIYEVPEELSGQTNHLSKNFPLTNRERSDPFINREVLANPKLP 480
433 RFGMDMETIGFAIVEPEELSGQTNHLSKNFPLTNRERSDPFINREVLANPKLP 492
481 EYIIVPSTFEPNKGDDFCIRVFSEKADYQAVDEIEANT--EEPDISEDDIDGVRLLFA 539
493 EYVVPSTFEPNKGDDFCIRVFSEKADYQAVDEIEANT--EEPDISEDDIDGVRLLFA 552
540 QLAEDDEIAFAELQTIIRRLAKRQDKSDGSIETCKIVMDLSDSGSKGLKEFYI 599
553 QLAEDDEIAFAELQTIIRRLAKRQDKSDGSIETCKIVMDLSDSGSKGLKEFYI 612
600 LMTIIOCKYKTYREIDVDRSGTMSYEMRKALKEAGFPMPCOLHOVIYARPADQLIIDF 659
613 LMTIIOCKYKTYREIDVDRSGTMSYEMRKALKEAGFPMPCOLHOVIYARPADQLIIDF 672
660 DNFRVCLVRLLETFKIFKQIDPENTGTIELDLISML 695
673 DNFRVCLVRLLETFKIFKQIDPENTGTIELDLISML 708

RESULT 9
US-08-835-099A-1
Sequence 1, Application US/08835099A
Patent No. 5874277
GENERAL INFORMATION:
APPLICANT: SHINTANI, Yasuehi
APPLICANT: NISHI, Kazuori

APPLICANT: KAWAMOTO, Tomohiro
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
FILE REFERENCE: AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,099A
FILING DATE: 04-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 083649/1996
FILING DATE: 05-APR-1996
APPLICATION NUMBER: 97105508.2
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 703 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-835-099A-1

Query Match 63.8%; Score 2370.5; DB 2; Length 703;
Best Local Similarity 61.5%; Pred. No. 9.2e-228;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

1 MAGIAKLADREAAAGLSHERAIKYLNDYALNCEIAETLRQDSFPALPSALGFE 60
1 MAAQAGVSRQARATQGLSNQNALKYLGDFFTLRQCLDSDVLFKDPFPAQPSALGY 60
61 KELGPSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWLLAIAASLTINEEL 120
61 KDLGPGSPQTOGIIMKRPTEICPSPOFIYAGATRTDICGALGDCWLLAIAASLTINEEL 120
121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFWMSAL 180
121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFWMSAL 180
121 LYRVPRDQDFQENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFWMSAL 180
181 LERAYAKINGCYALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILG 240
181 LERAYAKINGCYALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILG 240
181 LERAYAKINGCYALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILG 240
241 SIDTSAASEATLTKLVKGAHAYVTGAEEVNSGLQKLIIRPMWGEVMTGRMNCP 300
241 SIDTSAASEATLTKLVKGAHAYVTGAEEVNSGLQKLIIRPMWGEVMTGRMNCP 300
241 SIDTSAASEATLTKLVKGAHAYVTGAEEVNSGLQKLIIRPMWGEVMTGRMNCP 300
301 CPSMNTIDEEERELTRRHEDGEFWMSPDFLRHYSRLICNLTPDLTSDYKKWKLTK 360
301 APEWNHIDPRKRELDKVDGEFWMSPDFLRHYSRLICNLTPDLTSDYKKWKLTK 360
361 MGNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEB--DESGCTFLVGLIQHRRR 418
361 MGNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEB--DESGCTFLVGLIQHRRR 420

Matches	430;	Conservative	123;	Mismatches	143;	Indels	3;	Gaps	2
QY	1	MAGTAAATLADREAAEGSHERAIKYLANDYELARNHECLAGTLFQDPSPFAPLPSALGF	60						
Db	1	MAQAAGVSHSRAPATQGGISQNALKYLGQDFKTLRQCDLDSGVLKDPPEPPACSAALGY	60						
QY	61	KELAPYSKTRGMKMRPTETICADPOFIFIGATGATTDICQALGDCWMLAAALASLTNBEI	120						
Db	61	KDLGSPSPQIOGIIMKRPTELCPSPQPIVGGATTDIDCGGLGDCWMLAAALASLTNBEI	120						
QY	121	LARVPLNQSFOENYAGIFHFQFQWQYGEWEEVVVDRLPTQDGEILFVHASAGSEFWGAL	180						
Db	121	LYRVPLPDQPEQENYAGIFHFQFQWQYGEWEEVVVDRLPTQNGQLFLHSGQGNFWGAL	180						
QY	181	LEKRYAKINGCYEALSGATTGEGEDTGGIAENYELKCPPNIFKIIQKALQKSLIGC	240						
Db	181	LEKRYALINGCYEALSGATVEGEDTGGISEFYDLKKPPANNYQIIRKALCAGSLIGC	240						
QY	241	SIDTSAADSEAFIFQGLVKGHAASVYGAEEVSENGSLQKILRINPMGEVEMTGRMND	300						
Db	241	SIDVSSAAEAPATISQKLVKSHAASVYGEVFNQGHDEKILIRLNPMGEVEMSGAWSDD	300						
QY	301	CPSWNTIDPEBERELTFRHDEGEFWMSPFSLFRHYSRLICNTLPDITLSTYKKMKLTK	360						
Db	301	APENNHIDPFRKEELDKKVEDGEFWMELSDPVRFSLREICNLSPDLSSEYHKMLVL	360						
QY	361	MDGNMRGSGTAGGCRNYPNTFMNPOYLKLEEBEDEED--GSGGCTFLVGLIQKHRR	418						
Db	361	FNCHMTGTSIAGGCONPATYWTNPQFKIRLDEVEDDEBESIGBPCCVTLLGLMKNRW	420						
QY	419	QRKKGEMHTIGFRIYEVPELSQCTNHLKSNFLLNRAREBDTFINLBEVLRFKLP	478						
Db	421	RKRIGQGMELSIGAVVGVPRKELESHTAHGRDFPLAYQPSARSTYNALREVSGARLP	480						
QY	479	PGEYILVPTFEFPKODPFCIRVSEKKADYQAVDDEIENALFEEDISE--DIDIDGVRL	537						
Db	481	PGEILVPTSTFEPRKDSFCLIRVSEKKAAQALEIGDVVAGNPRYPHSEVUDQEDDQFRL	540						
QY	538	PAQIAGEDAEISAFELQTIILRRVLAKRODIKSDGFSIETCKINVMMLSDSGSKLKEF	597						
Db	541	PEKLAGDSEITANALKILINENASKRTDIKFDGPNINTGEMSLSDSNQGTIGAVEF	600						
QY	598	YILMTKIQKTYKRIEIDVDRSGTMSYENRKALBEAGFKMPQQLHQYTVARPADQULI	657						
Db	601	KTILMLKIQKLEYIEMEDYVNSGTIDAHENMTALRKAGFTLNSQVOCTIALRYACSKIGI	660						
QY	658	DEDNFVRCVLRLETLFKIFKQDPEENGTIELDLSMLC	696						
Db	661	NFDSFVACMIRLETLFLKFLSLDBDKGAWQLSLAEMLC	699						
RESULT 11									
US-08-835-099A-2									
Sequence 2, Application US/08835099A									
Patent No. 5874277									
GENERAL INFORMATION:									
APPLICANT: SHINTANI, Yasushi									
APPLICANT: NISHI, Kazuori									
APPLICANT: KAMAMOTO, Tomohiro									
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION									
NUMBER OF SEQUENCES: 18									
CORRESPONDENCE ADDRESS:									
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP									
STREET: 130 Water Street									
CITY: Boston									
STATE: MA									
COUNTRY: USA									
ZIP: 02109									
COMPUTER READABLE FORM:									
MEDIUM TYPE: Diskette									
COMPUTER: IBM Compatible									
OPERATING SYSTEM: DOS									

```

SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,099A
FILING DATE: 04-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 083649/1996
FILING DATE: 05-APR-1996
APPLICATION NUMBER: 97105508.2
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 712 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-835-099A-2

```

```

Query Match      63.8%; Score 2370.5; DB 2; Length 712;
Best Local Similarity 61.5%; Pred. No. 9.4e-228;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

```

```

QY 1 MAGIAIAKLAKDREAEGLGSHERAIKYLNODYEALRNECLEAGTLFODPSFPALPSALGF 60
DB 10 MAAQAGVSRGRATQGLSGNOMLKYLGODFKTLRQGLDGLVFXDPEFPACPSALGY 69
QY 61 KELGPISSKTRGMKRPTEICADPOITIGARTDTCQALGDCWLLAIAISLTINEE 120
DB 70 KDLPGSPQTOGIIWKRPTELCPSQPIVGGATRTDTCQGLGDCWLLAIAISLTINEE 129
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDGELFVHSAEGSEFMSAL 180
DB 130 LYRVVPRDPOENYAGIFHFQFQWYGEWVVDRLPTKDGELFVHSAEGSEFMSAL 189
QY 181 LEKAVAKINGCYEALSGATTEGFEPTGIAEWELKKPPNLFKIIQALOKGSLIGC 240
DB 190 LEKAVAKINGCYEALSGATTEGFEPTGIAEWELKKPPNLFKIIQALOKGSLIGC 249
QY 241 SIDITSAADSEALITFQKLVKGHANSVTGAEEVESNGSLQKILIRINPGEVEWTRMND 300
DB 250 SIDVSSAAEAALITSQKLVKSHAVSVTGVEEVNFQGHPEKILIRINPGEVEWTRMND 309
QY 301 CPSNNTIDPEERERTLRHDEGEFMSFSDFLRHYSRLICNLPTDILTSTYKKWKLTK 360
DB 310 APEWNNHIDPRRKEELDKVEDEGEFMSFSDFLRHYSRLICNLPTDILTSTYKKWKLTK 369
QY 361 MDGWRKSGTAGGCGNYPNTFMNPOYLKLEEBDEDEDD--GESGCFVLGLIQKHRR 418
DB 370 FNGHWTRSTAGGCGNYPATYNTNPOFKIRLDEVEDDEBESIGECCTVLGLQKHRR 429
QY 419 QKRGEDMHTTGFGLYEVPEELSGQTNILSKNFFLTNRABRSDTFLNREVLNRFKL 478
DB 430 KKRIGOGMLSTGYAVYQVPEKLESHHTDAHLSGDFLAVQPSARTSTYVNLREVSGRALP 489
QY 479 PGEYILVSTPEPNNDGFCIRVSEKADYQVAVDELEANEEDISE-DDIDGCVRL 537
DB 490 PGEYILVSTPEPNNDGFCIRVSEKADYQVAVDELEANEEDISE-DDIDGCVRL 549
QY 538 PAQIAGEDAEISAELEQTLIRVLAKRODIKSDGSFETCKIMVMDLSDSGKGLKEF 597
DB 550 FEKLAGKSEITANLKLINLNEABSKRDIKFDGNINITCREMISLSDSNGTGIAVEF 609
QY 598 YILMTKIQKYKIREIDVDRSGTWNSEYEMKALEAGFKMPCQLHQYIVARFADQILI 657
DB 610 KTLMLKIQKYLEIYETWTDVNHSGTIDAHENRILRAGFTLNSVOQOTIALRYACKLGI 669

```

```

DB 610 KTLMLKIQKYLEIYETWTDVNHSGTIDAHENRILRAGFTLNSVOQOTIALRYACKLGI 669
QY 658 DFDNFVACVRLTELEFKIFKQDPENTWTGLDLSMLC 696
DB 670 NFDSFVACVRLTELEFKIFKQDPENTWTGLDLSMLC 708

```

```

RESULT 12
US-09-157-349-2
Sequence 2, Application US/09157349
Patent No. 6068990
GENERAL INFORMATION:
APPLICANT: SHINTANI, Yasushi
APPLICANT: NISHI, Kazuo
APPLICANT: KAWAMOTO, Tomohiro
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
TITLE OF INVENTION: AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/157,349
FILING DATE:
FILING DATE:
APPLICATION NUMBER: 08/835,099
APPLICATION NUMBER: 97105508.2
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 712 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-157-349-2

```

```

Query Match      63.8%; Score 2370.5; DB 3; Length 712;
Best Local Similarity 61.5%; Pred. No. 9.4e-228;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

```

```

QY 1 MAGIAIAKLAKDREAEGLGSHERAIKYLNODYEALRNECLEAGTLFODPSFPALPSALGF 60
DB 10 MAAQAGVSRGRATQGLSGNOMLKYLGODFKTLRQGLDGLVFXDPEFPACPSALGY 69
QY 61 KELGPISSKTRGMKRPTEICADPOITIGARTDTCQALGDCWLLAIAISLTINEE 120
DB 70 KDLPGSPQTOGIIWKRPTELCPSQPIVGGATRTDTCQGLGDCWLLAIAISLTINEE 129
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDGELFVHSAEGSEFMSAL 180
DB 130 LYRVVPRDPOENYAGIFHFQFQWYGEWVVDRLPTKDGELFVHSAEGSEFMSAL 189
QY 181 LEKAVAKINGCYEALSGATTEGFEPTGIAEWELKKPPNLFKIIQALOKGSLIGC 240

```

Db 190 LERAYALNCGYELAGSGTVEGEDEFTGISEFYDKKRPANLYQIIRALCAGSLG 249
Qy 241 SIDIITSAADSEAITFOKLVKGHAYSTGAEEVNSGLOKLIIRNMGWEVMTGRANDN 300
Db 250 SIVVSAABEAITSOQLVKSHASTGVVEVNFQSGPEKILIRANMGWEVNSGASDD 309
Qy 301 CPSEMTIDPERERLTRRHDEGEFMSFSDFLRHYSLEICNLTPDTLSDTYKKWLT 360
Db 310 APEMNHIDPRKEELDKKVEDGEFMSLDFVRQFSRLICNLSPDLSSEVHKMNLVL 369
Qy 361 MDGNWRGSTAGGCRNYPNTFMNPOYLILKEDEDEBDG--ESGCTFLVGLIOKRRR 418
Db 370 FNGHWRGSTAGGCRNYPNTFMNPOYLILKEDEDEBDG--ESGCTFLVGLIOKRRR 429
Qy 419 ORKMGEDMTIGFIYVPEELSGQTNHLSKNPFLTNRARSDFINLREVLANRKL 478
Db 430 RKRIQGMSTIGAVYVQPELESHTDAHLGRDFLALYQPSARTSTYVNLREVSGRRLP 489
Qy 479 PGEYILVPSFTEPNKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 537
Db 490 PGEYILVPSFTEPNKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 549
Qy 538 PAOLAGDAISAELOTLIRVLAQRDIDSGFSIETCKINWMDSDSGKLGK 597
Db 550 FEKLAGDSEITANALKILNEAFSKQTDIKFDFNINTCREMISLSDSGTGTLGAVEF 609
Qy 598 YILMTKIQKYKIREIDVDRSGTMSYEMRKALBEAGFKMPCOLHQVIYVARFADOL 657
Db 610 KTMILKIQKLEIYEWEDYHSGTIDHNEHTALRKAGFTLNSGOQTILARVACSLGI 669
Qy 658 DFDNFVRLVLETLFKIFKQDPEBNTGTELDLSWLC 696
Db 670 NFDSPVACMIRLETLFKILFLDDEDKDQWQSLAEWLC 708

RESULT 13
US-09-422-869-27
Sequence 27, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOMISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BEIL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 703
TYPE: PRT
ORGANISM: RAT
US-09-422-869-27

Query Match 63.5%; Score 2356.5; DB 3; Length 703;
Best Local Similarity 60.9%; Pred. No. 2,3e-226;
Matches 427; Conservative 127; Mismatches 140; Indels 7; Gaps 3;
Qy 1 MAGIAAKLAKDRRAAGLSHERAIXLYNODYALNEBCLEAGTLFODSPFPAIPSLG 60
Db 1 MALIAAGVSKORAVAGLSGNQNAVKYLGODFETLRKQCLNSGVLFKDPFPACPSALG 60
Qy 61 KELGPYSSKTRGMWRKPTIEICADPOFIIGATRTDICGALGDCWLLAAIASLTINNEI 120

Db 61 KDLGPGSPDIOGIWAKRPTLCPNPQIFIVGATRTDIOGGLGDCWLLAAIASLTINNEI 120
Qy 121 LARVPLNOSFOBNVYIFHEQFMQGEVNVVDRLPRKDELLFVHSAEGSEFMSAL 180
Db 121 LRVPLPDOSFOQVAGIFHFQFMQGEVNVVDRLPRKQQLFVHSEEGSEFMSAL 180
Qy 181 LERAYAKINGCYEALSGATTEGEFEDFTGIAEMWELKPPNLFKILQALOGSLG 240
Db 181 LERAYALNCGYELAGSGTVEGEDEFTGISEFYDKKRPANLYQIIRALCAGSLG 249
Qy 241 SIDIITSAADSEAITFOKLVKGHAYSTGAEEVNSGLOKLIIRNMGWEVMTGRANDN 300
Db 241 SIVVSAABEAITSOQLVKSHASTGVVEVNFQSGPEKILIRANMGWEVNSGASDD 309
Qy 301 CPSEMTIDPERERLTRRHDEGEFMSFSDFLRHYSLEICNLTPDTLSDTYKKWLT 360
Db 301 APEMNHIDPRKEELDKKVEDGEFMSLDFVRQFSRLICNLSPDLSSEVHKMNLVL 369
Qy 361 MDGNWRGSTAGGCRNYPNTFMNPOYLILKEDEDEBDG--ESGCTFLVGLIOKRRR 418
Db 361 FNGHWRGSTAGGCRNYPNTFMNPOYLILKEDEDEBDG--ESGCTFLVGLIOKRRR 429
Qy 419 ORKMGEDMTIGFIYVPEELSGQTNHLSKNPFLTNRARSDFINLREVLANRKL 478
Db 421 ORKMGEDMTIGFIYVPEELSGQTNHLSKNPFLTNRARSDFINLREVLANRKL 480
Qy 479 PGEYILVPSFTEPNKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 537
Db 481 PGEYILVPSFTEPNKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 549
Qy 538 PAOLAGDAISAELOTLIRVLAQRDIDSGFSIETCKINWMDSDSGKLGK 597
Db 539 SLEEFYAGKSEISANQKRVNLVSKRDMFDFGNITCREMISLSDSGTGTLGAVEF 609
Qy 598 YILMTKIQKYKIREIDVDRSGTMSYEMRKALBEAGFKMPCOLHQVIYVARFADOL 657
Db 599 EFKTLMKIRTYLEIFQEMDHNVGTIEAHEMTALKAGFTLNSGOQTILARVACSLGI 669
Qy 656 IIDPNDVRLVLETLFKIFKQDPEBNTGTELDLSWLC 696
Db 659 GVDNPFVACMIRLETLFKILFLDDEDKDQWQSLAEWLC 708

RESULT 14
US-09-308-345A-49
Sequence 49, Application US/09308345A
Patent No. 6569665
GENERAL INFORMATION:
APPLICANT: BOEHM, Thomas;
APPLICANT: DEAR, Neil T.
TITLE OF INVENTION: No. 6569665el calpains, their preparation and use
FILE REFERENCE: 0050/4/576
CURRENT APPLICATION NUMBER: US/09/308,345A
CURRENT FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: WordPerfect v. 6.1
SEQ ID NO 49
LENGTH: 703
TYPE: PRT
ORGANISM: rat
US-09-308-345A-49

Query Match 62.2%; Score 2308.5; DB 4; Length 703;
Best Local Similarity 60.1%; Pred. No. 1,5e-221;
Matches 421; Conservative 125; Mismatches 148; Indels 7; Gaps 3;
Qy 1 MAGIAAKLAKDRRAAGLSHERAIXLYNODYALNEBCLEAGTLFODSPFPAIPSLG 60
Db 1 MALIAAGVSKORAVAGLSGNQNAVKYLGODFETLRKQCLNSGVLFKDPFPACPSALG 60
Qy 61 KELGPYSSKTRGMWRKPTIEICADPOFIIGATRTDICGALGDCWLLAAIASLTINNEI 120

```

Db      61 KDLPGSPDIOGIWVKRPTLCPNPQFIVGATRTDROGGLVDCWLLAALASITLNEKL 120
Qy      121 LARVPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFMSAL 180
Db      121 LARVPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFMSAL 180
Qy      181 LEKAYKINGCYEALSGGATTEGFEDFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
Db      181 LEKAYKINGCYEALSGGATTEGFEDFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
Qy      241 SIDTSAADSEAITFOKLYKGAHAYSVTGAEVEESNGSLQKIRIRNPMGEVMTGRANDN 300
Db      241 SIDTSAADSEAITFOKLYKGAHAYSVTGAEVEESNGSLQKIRIRNPMGEVMTGRANDN 300
Qy      301 CPSNWTIDPEERERLTRRHEDEEFMMSFSDFLRHSRLICNLTPDITLSDTYKKWLTJK 360
Db      301 CPSNWTIDPEERERLTRRHEDEEFMMSFSDFLRHSRLICNLTPDITLSDTYKKWLTJK 360
Qy      361 MDGNWRGSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--ESGCTFLVGLIQKHRR 418
Db      361 MDGNWRGSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--ESGCTFLVGLIQKHRR 418
Qy      419 QKRGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRARERDPTFNLREYNRFLP 478
Db      419 QKRGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRARERDPTFNLREYNRFLP 478
Qy      479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 535
Db      479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 535
Qy      536 RLPAQLAGEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGSGKLGLK 595
Db      536 RLPAQLAGEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGSGKLGLK 595
Qy      596 EFLYLMKIRTYLEIFQEMDNHVGITIEAHHRKALKKAGPTLNNQVOOTIAMRYASCKL 658
Db      596 EFLYLMKIRTYLEIFQEMDNHVGITIEAHHRKALKKAGPTLNNQVOOTIAMRYASCKL 658
Qy      656 IIDFNFVRLVLETLFKIFKQDPENGTITLIDLSMLC 696
Db      656 IIDFNFVRLVLETLFKIFKQDPENGTITLIDLSMLC 696
Qy      697 GVDNPGFVACMIRLETLFKLFRLLDKQNGIYQSLAEMLC 699
Db      697 GVDNPGFVACMIRLETLFKLFRLLDKQNGIYQSLAEMLC 699

RESULT 15
US-09-653-839-2
; Sequence 2, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
; FILE REFERENCE: LEX-0038-USA
; CURRENT APPLICATION NUMBER: US/09/653,839
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/152,057
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 739
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-653-839-2

```

Query Match 54.4%; Score 2020; DB 4; Length 739;
 Best Local Similarity 52.5%; Pred. No. 1,1e-192;
 Matches 366; Conservative 134; Mismatches 193; Indels 4; Gaps 2;

```

Qy      3 GIAKLAKORAEAEGLGSHERRAIKYLNOYEAALRNECLEAGTLFODSPFALSALGFKE 62
Db      37 GMAVHINNSLKKAKGVQGHQDNMQNFGNQSFEEELRAACLRKGELEFEDLPFAEPSSISGFKD 96
Qy      63 LGFYSSKTRGMKRRKPRLEICADPOFIIIGATRTDIOGALGDCWLLAALASITLNEELIA 122
Db      97 LGFNSKNVQNIISWQRPDIINPLFTMDGISPFIDIOGILGDCWLLAALASITLCPPLLY 156
Qy      123 RVPPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFMSAL 182
Db      157 RVPPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAEGSEFMSAL 182
Qy      183 KAYAKINGCYEALSGGATTEGFEDFTGIAEWELKPPNLFKIIQKALQKSLGCG 242
Db      217 KAYAKINGCYEALSGGATTEGFEDFTGIAEWELKPPNLFKIIQKALQKSLGCG 242
Qy      243 DITSADSEAITFOKLYKGAHAYSVTGAEVEESNGSLQKIRIRNPMGEVMTGRANDN 302
Db      277 EVTSDELEEMTKMLVRGHAYSVTGLQDVHVRGKMETLIRVNPWGRLEMNAGMSDSAR 336
Qy      303 SMTIDPEERERLTRRHEDEEFMMSFSDFLRHSRLICNLTPDITLSDTYKKWLTJK 362
Db      337 EMBEVASDIOMQLHKTEDDEFMMSYODFLNFTLLEICNLTPDITLSDTYKKWLTJK 362
Qy      363 GNNRSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--CTFLVGLIQKHRRQ 419
Db      397 GNNRSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--CTFLVGLIQKHRRQ 419
Qy      420 RKNGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRARERDPTFNLREYNRFLP 479
Db      457 RKNGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRARERDPTFNLREYNRFLP 479
Qy      480 GEYILVSTPEPNKDGFCIRVFSEKKADYQAVDD--EIEANLEPDISDDIDGVR 538
Db      517 GEYILVSTPEPNKDGFCIRVFSEKKADYQAVDD--EIEANLEPDISDDIDGVR 538
Qy      539 AOLAGEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGSGKLGLKEFY 598
Db      577 AOLAGEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGSGKLGLKEFY 598
Qy      599 ILMTKIOKTYREIDVDSGTMNSYEMKALBEAGFKRPPCQLOHYIVARPDQOL 658
Db      637 ILMTKIOKTYREIDVDSGTMNSYEMKALBEAGFKRPPCQLOHYIVARPDQOL 658
Qy      659 FDNFVRLVLETLFKIFKQDPENGTITLIDLSMLC 695
Db      697 FDNFVRLVLETLFKIFKQDPENGTITLIDLSMLC 695

```

Search completed: July 24, 2003, 12:58:04
 Job time : 32.9606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:27:29 ; Search time 89.9793 Seconds
(without alignments)
1234.824 Million cell updates/sec

Title: us-09-884-319a-7

Perfect score: 3713
Sequence: 1 MAGIAAKLANDREAREGLGS.....PENTGTETLDLISWLCFSVL 700

Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_190jun03.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	AAW19992	Human CAMP used to
2	3713	100.0	700	AAW19992	Human CAMP used to
3	3713	100.0	700	AAW19992	Human CAMP used to
4	3694	99.5	700	AAW19992	Human CAMP used to
5	3518	94.7	700	AAW19992	Human CAMP used to
6	2435	65.6	713	AAW19992	Human CAMP used to
7	2370.5	63.8	703	AAW19992	Human CAMP used to
8	2370.5	63.8	703	AAW19992	Human CAMP used to
9	2367.5	63.8	703	AAW19992	Human CAMP used to

10	2367.5	63.8	703	23	AAW19992	Human CAMP used to
11	2020	54.4	703	22	AAW19992	Human CAMP used to
12	2014	54.2	702	22	AAW19992	Human CAMP used to
13	2014	54.2	702	22	AAW19992	Human CAMP used to
14	2014	54.2	702	22	AAW19992	Human CAMP used to
15	2014	54.2	702	22	AAW19992	Human CAMP used to
16	2012	54.2	702	21	AAW19992	Human CAMP used to
17	2002	53.9	702	21	AAW19992	Human CAMP used to
18	2002	53.9	702	21	AAW19992	Human CAMP used to
19	1967	53.0	703	22	AAW19992	Human CAMP used to
20	1961	52.8	686	22	AAW19992	Human CAMP used to
21	1925.5	51.9	707	20	AAW19992	Human CAMP used to
22	1889.5	50.9	703	24	AAW19992	Human CAMP used to
23	1888.5	50.7	767	23	AAW19992	Human CAMP used to
24	1882	50.7	690	24	AAW19992	Human CAMP used to
25	1862.5	50.2	821	17	AAW19992	Human CAMP used to
26	1610.5	43.4	791	22	AAW19992	Human CAMP used to
27	1594.5	42.9	828	22	AAW19992	Human CAMP used to
28	1532.5	41.3	1069	22	AAW19992	Human CAMP used to
29	1495.5	40.3	399	22	AAW19992	Human CAMP used to
30	1453	39.5	720	23	AAW19992	Human CAMP used to
31	1354	36.5	565	22	AAW19992	Human CAMP used to
32	1335	36.0	708	23	AAW19992	Human CAMP used to
33	1335	36.0	708	23	AAW19992	Human CAMP used to
34	1328.5	35.8	648	23	AAW19992	Human CAMP used to
35	1287	34.7	251	18	AAW19992	Human CAMP used to
36	1287	34.7	251	18	AAW19992	Human CAMP used to
37	1218	32.8	300	22	AAW19992	Human CAMP used to
38	1150	31.0	687	24	AAW19992	Human CAMP used to
39	1148.5	30.9	694	24	AAW19992	Human CAMP used to
40	1134	30.5	711	23	AAW19992	Human CAMP used to
41	1126.5	30.3	343	22	AAW19992	Human CAMP used to
42	1120.5	30.2	462	23	AAW19992	Human CAMP used to
43	1120.5	30.2	518	23	AAW19992	Human CAMP used to
44	1111.5	29.9	659	24	AAW19992	Human CAMP used to
45	1110	29.9	656	24	AAW19992	Human CAMP used to

ALIGNMENTS

RESULT 1
AAW19992 standard; Protein; 700 AA.
ID AAW19992 standard; Protein; 700 AA.
AC AAW19992;
DT 27-AUG-1997 (first entry)
XX
DE Human CAMP used to identify inhibitors of interleukin-1 activity.
XX
KW IL-1; interleukin; receptor; ligand; screening assay; inhibitor;
KW IL-1 mediated response; inflammation; inflammatory; antibody;
KW intracellular domain; CAMP; calcium activated neutral protease.
XX
OS Homo sapiens.
XX
PN W09640907-A1.
PD 19-DEC-1996.
PF 06-MAY-1996; 96MO-US06363.
PR 07-JUN-1995; 95US-0487942.
XX
XX (GENY) GENETICS INST INC.
XX Graham J, Lin L;
XX WPI; 1997-052315/05.
XX
XX Interleukin-1 receptor intracellular ligand proteins and related DNA
PT - used to identify inhibitors of the proteins for treatment of

PT inflammation
 XX Claim 14; Page 36-38; 54pp. English.
 XX
 CC AAM19992 represents human calcium activated neutral protease (CANP).
 CC This protein was found to have an area of high homology with an
 CC interleukin-1 receptor (IL-1-R) intracellular ligand (encoded by cDNA
 CC clone 14w, see AAT71218) and thus will display some of the same
 CC properties of this protein. IL-1-R intracellular ligand proteins are
 CC used to screen for agents (e.g. antibodies) that are capable of
 CC inhibiting or blocking the binding of an IL-1-R intracellular ligand
 CC to the intracellular domain of IL-1-R, i.e. inhibitors of IL-1
 CC activity. Such agents can be used to treat inflammatory conditions.
 XX
 SQ Sequence 700 AA;
 Query Match 100.0%; Score 3713; DB 18; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGIAAKLADREAEGLGSHERAIKYLNODYEALNCELEAGTLFODSPFPALPSALGF 60
 DB 1 MAGIAAKLADREAEGLGSHERAIKYLNODYEALNCELEAGTLFODSPFPALPSALGF 60
 QY 61 KELGPYSSKTRGMKRWKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTNBEI 120
 DB 61 KELGPYSSKTRGMKRWKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTNBEI 120
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWELKRPPLFKIIQKALQKSLG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWELKRPPLFKIIQKALQKSLG 240
 QY 241 SIDITSAADSEAITFOKLKVGAVSVTGAEEVSENGSLQKILIRINPGEVEMTGRWNN 300
 DB 241 SIDITSAADSEAITFOKLKVGAVSVTGAEEVSENGSLQKILIRINPGEVEMTGRWNN 300
 QY 301 CPSMWTIDPERERELTRHEDGEEFMSFSDLRHYSRLICNLTPDITLSTYKKMKLTK 360
 DB 301 CPSMWTIDPERERELTRHEDGEEFMSFSDLRHYSRLICNLTPDITLSTYKKMKLTK 360
 QY 361 MDGMMRGSTAGGCRNYPNTFEMNPOYLKLEBEDDEDEDESGCTFLVGLIOKRRROR 420
 DB 361 MDGMMRGSTAGGCRNYPNTFEMNPOYLKLEBEDDEDEDESGCTFLVGLIOKRRROR 420
 QY 421 KMGEDMHTIGFGIYEVPEELSQQTNILSKNPFLLNRAERSDFTINLREVLNRPKLP 480
 DB 421 KMGEDMHTIGFGIYEVPEELSQQTNILSKNPFLLNRAERSDFTINLREVLNRPKLP 480
 QY 481 EYILVPSFEPENKGDPCIRVFSEKADYQAVDEIEANLESPDISEDDIDGVRRLFAQ 540
 DB 481 EYILVPSFEPENKGDPCIRVFSEKADYQAVDEIEANLESPDISEDDIDGVRRLFAQ 540
 QY 541 LAGSDAISIAPFELQTLIRRYLAKRODIDSQFSIETCKIMWMDLSDSGKLGLKEPITL 600
 DB 541 LAGSDAISIAPFELQTLIRRYLAKRODIDSQFSIETCKIMWMDLSDSGKLGLKEPITL 600
 QY 601 WTKIQTKYIREIDVDSGTMNSYEMRKALEAGFKAPCOLHQYAVAFADQULIIDFD 660
 DB 601 WTKIQTKYIREIDVDSGTMNSYEMRKALEAGFKAPCOLHQYAVAFADQULIIDFD 660
 QY 661 NFVACLVRLFTLFIKQKQDPENTGTIFLDLISWLCFSVL 700
 DB 661 NFVACLVRLFTLFIKQKQDPENTGTIFLDLISWLCFSVL 700
 QY 661 NFVACLVRLFTLFIKQKQDPENTGTIFLDLISWLCFSVL 700
 DB 661 NFVACLVRLFTLFIKQKQDPENTGTIFLDLISWLCFSVL 700

RESULT 2
 AAB37797
 ID AAB37797 standard; Protein; 700 AA.

AC AAB37797;
 XX 23-FEB-2001 (first entry)
 DT
 XX
 DE Human interleukin-1 receptor intracellular ligand protein #4.
 KW Human; interleukin-1; IL-1; IL-1alpha; IL-1beta; IL-1 receptor;
 KW antiinflammatory; haemostatic; antibacterial; immunosuppressive;
 KW immunomodulatory; cardiac; cytostatic; neuroprotective; respiratory;
 KW inflammation; infection; sepsis; cachexia; autoimmune disorder;
 KW cardiovascular disorder; chronic myelogenous leukaemia;
 KW multiple sclerosis; inflammatory bowel disease; Crohn's disease.
 XX
 OS Homo sapiens.
 PN W0200064479-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 26-APR-2000; 2000MO-US11700.
 XX
 PR 27-APR-1999; 99US-0301274.
 XX
 PA (ANTI-) ANTIBODY SYSTEMS INC.
 XX
 PI Fredeking TM, Ignatyev GM;
 XX
 DR WPI, 2000-679646/66.
 XX
 PT Novel compositions comprising tetracycline or tetracycline-like
 PT compounds for the treatment and/or prevention of acute inflammatory
 PT responses and diseases, e.g. septic shock and immune complex-induced
 PT colitis -
 XX
 PS Disclosure; Page 159-162; 183pp; English.
 XX
 CC The present sequence is given in a specification relating to novel
 CC compositions and methods containing tetracycline or tetracycline-like
 CC compounds for treating and/or preventing acute inflammatory responses and
 CC diseases. Such diseases include acute inflammatory conditions associated
 CC with viral haemorrhagic diseases (including diseases caused by
 CC Bunyaviridae, Filoviridae, Flaviviridae or Arenaviridae viruses),
 CC parasitic diseases, bacterial infections, sepsis, cachexia, autoimmune
 CC disorders, acute cardiovascular events, chronic myelogenous leukaemia and
 CC transplanted bone marrow-induced graft-versus-host disease, septic shock,
 CC immune complex-induced colitis, cerebrospinal fluid inflammation,
 CC multiple sclerosis, inflammatory responses associated with trauma,
 CC systemic inflammatory response syndrome (SIRS), adult respiratory
 CC distress syndrome (ARDS), acute liver failure, inflammatory bowel disease
 CC and Crohn's disease.
 XX
 SO Sequence 700 AA;
 Query Match 100.0%; Score 3713; DB 21; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGIAAKLADREAEGLGSHERAIKYLNODYEALNCELEAGTLFODSPFPALPSALGF 60
 DB 1 MAGIAAKLADREAEGLGSHERAIKYLNODYEALNCELEAGTLFODSPFPALPSALGF 60
 QY 61 KELGPYSSKTRGMKRWKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTNBEI 120
 DB 61 KELGPYSSKTRGMKRWKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTNBEI 120
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVWVDDRLPTKDGELLFVHSAEGSEFWAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWELKRPPLFKIIQKALQKSLG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEWELKRPPLFKIIQKALQKSLG 240

QY 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVSNGLQKILIRNPWGEVWNTGWMND 300
 Db 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVSNGLQKILIRNPWGEVWNTGWMND 300
 QY 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFDFLRHYSLEICNLTPTLTSDTYKKMKLTK 360
 Db 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFDFLRHYSLEICNLTPTLTSDTYKKMKLTK 360
 QY 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDESGCTFLVGLIQHRRRQ 420
 Db 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDESGCTFLVGLIQHRRRQ 420
 QY 421 KMGEDMHTTIGFGIYEVPEELSGQTNILHSKNPFLTNRABRSPTFLNREVLNRPFLPG 480
 Db 421 KMGEDMHTTIGFGIYEVPEELSGQTNILHSKNPFLTNRABRSPTFLNREVLNRPFLPG 480
 QY 481 EYLLVSTPEPNMGDFCTRVSEKKADYQAVDDEIENALEEFDISEDDIDGVRRLFAQ 540
 Db 481 EYLLVSTPEPNMGDFCTRVSEKKADYQAVDDEIENALEEFDISEDDIDGVRRLFAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGLKEFYIL 600
 Db 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGLKEFYIL 600
 QY 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPCQLHQVIVARPADQLIIDFD 660
 Db 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPCQLHQVIVARPADQLIIDFD 660
 QY 661 NFPRCLVRLETLFKIFKQDPENTGTIELDLISMLCFSYL 700
 Db 661 NFPRCLVRLETLFKIFKQDPENTGTIELDLISMLCFSYL 700

RESULT 3

AAE25059 standard; Protein; 700 AA.

XX AAE25059;
 XX 30-OCT-2002 (first entry)
 XX Human calpain protein #2.
 XX Human; calpain; nervous system disorder; amyotrophic lateral sclerosis;
 KW Parkinson's disease; dementia; genito-urinary system disorder; stroke;
 KW Alzheimer's disease; multiple sclerosis; benign prostate hyperplasia;
 KW urinary incontinence; gene therapy; cyostatic; nootropic; utropatic;
 KW neuroprotective.

XX Homo sapiens.
 XX PN W0200248326-A2.
 XX 20-JUN-2002.

XX 14-DEC-2001; 2001WO-EP14819.
 XX 14-DEC-2000; 2000US-255058P.

XX (FARB) BAYER AG.
 XX Ramakrishnan S;
 XX WPI; 2002-537625/57.

XX New human calpain polypeptide, useful for treating peripheral and
 PT central nervous system disorder and genito-urinary system disorders
 PT including urinary incontinence and benign prostate hyperplasia -
 XX Disclosure; Page 101-104; 110pp; English.

XX The invention relates to novel human calpain proteins and polynucleotides
 CC encoding such proteins. Calpain sequences of the invention are useful for

CC treating, ameliorating or correcting dysfunctions or diseases such as
 CC peripheral or central nervous system (CNS) disorders (e.g., Parkinson's
 CC disease, Alzheimer's disease, multiple sclerosis, stroke, amyotrophic
 CC lateral sclerosis, dementia) and genito-urinary system disorders such
 CC as urinary incontinence and benign prostate hyperplasia. They are also
 CC used in gene therapy. The present sequence is human calpain protein.

Sequence 700 AA;

Query Match 100.0%; Score 3713; DB 23; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIAAKLAKDRRAEAGLSHERAIKYLNQDYEALENECELAGTLPQDSFPAPALGF 60
 Db 1 MGIAAKLAKDRRAEAGLSHERAIKYLNQDYEALENECELAGTLPQDSFPAPALGF 60
 QY 61 KELGPPYSSTKRGKWRKRPTEICADPOFITGATRTDICGALGDCWLLAIASTLTNEET 120
 Db 61 KELGPPYSSTKRGKWRKRPTEICADPOFITGATRTDICGALGDCWLLAIASTLTNEET 120
 QY 121 LARVVPNLNOSFOENVAGIFHPQWQYGEWEVVVDRLPTKQCELLFVHSABGSEFMSAL 180
 Db 121 LARVVPNLNOSFOENVAGIFHPQWQYGEWEVVVDRLPTKQCELLFVHSABGSEFMSAL 180
 QY 181 LEKAVAKINGCYEALSGCATTEGFEEDFTGIAEWYELAKKPPNLFKIIQALOKSLLGC 240
 Db 181 LEKAVAKINGCYEALSGCATTEGFEEDFTGIAEWYELAKKPPNLFKIIQALOKSLLGC 240
 QY 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVSNGLQKILIRNPWGEVWNTGWMND 300
 Db 241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVSNGLQKILIRNPWGEVWNTGWMND 300
 QY 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFDFLRHYSLEICNLTPTLTSDTYKKMKLTK 360
 Db 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFDFLRHYSLEICNLTPTLTSDTYKKMKLTK 360
 QY 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDESGCTFLVGLIQHRRRQ 420
 Db 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDESGCTFLVGLIQHRRRQ 420
 QY 421 KMGEDMHTTIGFGIYEVPEELSGQTNILHSKNPFLTNRABRSPTFLNREVLNRPFLPG 480
 Db 421 KMGEDMHTTIGFGIYEVPEELSGQTNILHSKNPFLTNRABRSPTFLNREVLNRPFLPG 480
 QY 481 EYLLVSTPEPNMGDFCTRVSEKKADYQAVDDEIENALEEFDISEDDIDGVRRLFAQ 540
 Db 481 EYLLVSTPEPNMGDFCTRVSEKKADYQAVDDEIENALEEFDISEDDIDGVRRLFAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGLKEFYIL 600
 Db 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGLKEFYIL 600
 QY 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPCQLHQVIVARPADQLIIDFD 660
 Db 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPCQLHQVIVARPADQLIIDFD 660
 QY 661 NFPRCLVRLETLFKIFKQDPENTGTIELDLISMLCFSYL 700
 Db 661 NFPRCLVRLETLFKIFKQDPENTGTIELDLISMLCFSYL 700

RESULT 4

AAB86128 standard; protein; 700 AA.

XX AAB86128;
 XX 27-JUL-2001 (first entry)

XX Human calpain 80kDa subdomain protein fragment.

XX Calpain; calcium-activated cysteine proteinase; human; spatial structure;

KM Ca-activated cysteine proteinase; protein coordinate data; treatment;
KM structure-function study; ischemic condition; muscular dystrophy; tumor;
KM muscular; antitumor.
XX
OS Homo sapiens.
XX
PN EPI108779-A2.
XX
PD 20-JUN-2001.
XX
PF 13-DEC-2000; 2000EP-0127369.
XX
PR 14-DEC-1999; 99DE-1060225.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (PROT-) PROTEROS BIOSTRUCTURES GMBH.
XX
PI Strobl S, Fernandez-Catalan C, Bode W, Huber R, Suzuki K;
XX
XX WPI; 2001-376928/40.
XX
PT Spatial structures containing calpain-derived polypeptides, useful for
PT identifying calpain modulators and substrates, potentially useful e.g.
PT as antitumor agents -
XX
XX Claim 15; Fig 4; 182pp; German.
XX
CC This invention describes the novel spatial structure of human and rat
CC neutral calcium-activated cysteine protease (calpain) family. The spatial
CC structure (especially crystalline forms) are used for structure-function
CC studies, particularly for identifying (pseudo)substrates, inhibitors and
CC activators of calpains, potentially useful for treatment of ischemic
CC conditions, muscular dystrophy and/or tumors. The products of the
CC invention have anti-ischemic, muscular and antitumor activity. This
CC sequence represents the human calpain 80kDa subunit described in the
CC method of the invention.
XX
SQ Sequence 700 AA;

Query Match 99.5%; Score 3694; DB 22; Length 700;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 696; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKDREAEGLSHERAIKYLNDYELRNCEAGTLFQDPSPFPAISALGF 60
DB 1 MAGIAAKLAKDREAEGLSHERAIKYLNDYELRNCEAGTLFQDPSPFPAISALGF 60
QY 61 KELGPYSKTRGMWRKRPTEICADPQITIGATRTDICOALGDCWLAIAIASLTNEEI 120
DB 61 KELGPYSKTRGMWRKRPTEICADPQITIGATRTDICOALGDCWLAIAIASLTNEEI 120
QY 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFWSAL 180
DB 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFWSAL 180
QY 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFWSAL 180
DB 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFWSAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
DB 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
QY 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
DB 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
QY 241 SIDITSADSAITFOKLKVGHAIVTGAEEVNSGLOKILIRINPGEVETGRWMDN 300
DB 241 SIDITSADSAITFOKLKVGHAIVTGAEEVNSGLOKILIRINPGEVETGRWMDN 300
QY 241 SIDITSADSAITFOKLKVGHAIVTGAEEVNSGLOKILIRINPGEVETGRWMDN 300
DB 241 SIDITSADSAITFOKLKVGHAIVTGAEEVNSGLOKILIRINPGEVETGRWMDN 300
QY 301 CPSNNTTIDPEERELTRRHEDGFEFWSFSDPLRHYSRLICNLTPDITLSTTYKKKLT 360
DB 301 CPSNNTTIDPEERELTRRHEDGFEFWSFSDPLRHYSRLICNLTPDITLSTTYKKKLT 360
QY 301 CPSNNTTIDPEERELTRRHEDGFEFWSFSDPLRHYSRLICNLTPDITLSTTYKKKLT 360
DB 301 CPSNNTTIDPEERELTRRHEDGFEFWSFSDPLRHYSRLICNLTPDITLSTTYKKKLT 360
QY 361 MDGWMRSGTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTPLVGLIOKHRROR 420
DB 361 MDGWMRSGTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTPLVGLIOKHRROR 420
QY 361 MDGWMRSGTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTPLVGLIOKHRROR 420
DB 361 MDGWMRSGTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTPLVGLIOKHRROR 420
QY 421 KMGEDMTTIGFGLIYVPEELSQQTNILSKNFFLTNRARSDPTFINIREVLNPKLP 480
DB 421 KMGEDMTTIGFGLIYVPEELSQQTNILSKNFFLTNRARSDPTFINIREVLNPKLP 480

DB 421 KMGEDMTTIGFGLIYVPEELSQQTNILSKNFFLTNRARSDPTFINIREVLNPKLP 480
QY 481 EYILVPESTPEPNKDGPCIRVFSEKKADYQAVDEIEANLEPDISIDDGVRRLEAQ 540
DB 481 EYILVPESTPEPNKDGPCIRVFSEKKADYQAVDEIEANLEPDISIDDGVRRLEAQ 540
QY 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
DB 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
QY 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALFEAGFMPQQLHOVYARPADOLIIDPD 660
DB 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALFEAGFMPQQLHOVYARPADOLIIDPD 660
QY 661 NFVRCVRLTETLFKIFKQIDPENTGITEIDLISMLGFSYL 700
DB 661 NFVRCVRLTETLFKIFKQIDPENTGITEIDLISMLGFSYL 700

RESULT 5
AAB86130 ID AAB86130 standard; protein; 700 AA.
XX
XX AAB86130;
XX
XX 27-JUL-2001 (first entry)
XX
DE Rat calpain 80kDa subdomain protein fragment.
XX
KM Calpain; calcium-activated cysteine proteinase; human; spatial structure;
KM Ca-activated cysteine proteinase; protein coordinate data; treatment;
KM structure-function study; ischemic condition; muscular dystrophy; tumor;
KM muscular; antitumor.
XX
OS Rattus norvegicus.
XX
PN EPI108779-A2.
XX
PD 20-JUN-2001.
XX
PF 13-DEC-2000; 2000EP-0127369.
XX
PR 14-DEC-1999; 99DE-1060225.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (PROT-) PROTEROS BIOSTRUCTURES GMBH.
XX
PI Strobl S, Fernandez-Catalan C, Bode W, Huber R, Suzuki K;
XX
XX WPI; 2001-376928/40.
XX
PT Spatial structures containing calpain-derived polypeptides, useful for
PT identifying calpain modulators and substrates, potentially useful e.g.
PT as antitumor agents -
XX
XX Claim 15; Fig 6; 182pp; German.
XX
CC This invention describes the novel spatial structure of human and rat
CC neutral calcium-activated cysteine protease (calpain) family. The spatial
CC structure (especially crystalline forms) are used for structure-function
CC studies, particularly for identifying (pseudo)substrates, inhibitors and
CC activators of calpains, potentially useful for treatment of ischemic
CC conditions, muscular dystrophy and/or tumors. The products of the
CC invention have anti-ischemic, muscular and antitumor activity. This
CC sequence represents the rat calpain 80kDa subunit described in the
CC method of the invention.
XX
SQ Sequence 700 AA;

Query Match 94.7%; Score 3518; DB 22; Length 700;
Best Local Similarity 93.4%; Pred. No. 7.7e-310;
Matches 654; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

```

QY 1 MAGIATLADREAEGLGSHERAIKYLNDYELRNCELEAGTLFODPSFPAIPSLG 60
DB 1 MAGIATLADREAEGLGSHERAIKYLNDYELRNCELEAGTLFODPSFPAIPSLG 60
QY 61 KELGPPSSKTRGMRKWPTEICADPOITIGGATTTDCCGALGDCWLLAIASTLTNEE 120
DB 61 KELGPPSSKTRGMRKWPTEICADPOITIGGATTTDCCGALGDCWLLAIASTLTNEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDRLPTKDGELFVHSAEGSEFMSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDRLPTKDGELFVHSAEGSEFMSAL 180
QY 181 LERAYAKINGCYEALSGATTEGPDPTGGIAEWELKKPPNLFKTIQALOKGSLG 240
DB 181 LERAYAKINGCYEALSGATTEGPDPTGGIAEWELKKPPNLFKTIQALOKGSLG 240
QY 241 SIIITSAADSEALTFQGLVGHAYSVTGAEEVSSGLQKILIRNPMGQVEMTGKWNND 300
DB 241 SIIITSAADSEALTFQGLVGHAYSVTGAEEVSSGLQKILIRNPMGQVEMTGKWNND 300
QY 301 CPSPMTIDPEERERLTERHDEGEFMSFSDFLRHSRLCINLTPDTLSDTYKKWLT 360
DB 301 CPSPMTIDPEERERLTERHDEGEFMSFSDFLRHSRLCINLTPDTLSDTYKKWLT 360
QY 361 MDGWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDEGSGCTFLVGLIQHRRR 420
DB 361 MDGWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDEGSGCTFLVGLIQHRRR 420
QY 421 KMGEDHTTIGFGEVEEELSGOTNHLNKNPFLTNARSDPTNLRVLRFKLP 480
DB 421 KMGEDHTTIGFGEVEEELSGOTNHLNKNPFLTNARSDPTNLRVLRFKLP 480
QY 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
DB 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEPDISDDIDGVRRLFAQ 540
QY 541 LAGBDAISAFELQTIIRVLAARODIKSDGFSIETCKIWMDLSDSGSLGKKEFY 600
DB 541 LAGBDAISAFELQTIIRVLAARODIKSDGFSIETCKIWMDLSDSGSLGKKEFY 600
QY 601 WTKIOKQKTYRELDVDSGTMSYEMRKALEAGFMPQOLHOVYARADQOLITDP 660
DB 601 WTKIOKQKTYRELDVDSGTMSYEMRKALEAGFMPQOLHOVYARADQOLITDP 660
QY 661 NFVRCVRLTEFLFKIFKQDPENTGTLELDLSWLCFSVL 700
DB 661 NFVRCVRLTEFLFKIFKQDPENTGTLELDLSWLCFSVL 700

```

```

RESULT 6
AAE23085
ID AAE23085 standard; Protein: 713 AA.
XX
AC AAE23085;
XX
DT 21-AUG-2002 (first entry)
XX
DE Calcium-activated neutral protease protein.
XX
KW Transgenic; transgenic animal; pharmacological therapy; gene therapy;
KW phenotype modulation; calcium-activated neutral protease; enzyme;
KW genetic disease.
XX
OS Unidentified.
XX
PN MO200203787-A2.
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21427.
XX
PR 06-JUL-2000; 2000US-216109P.
PR 06-JUL-2000; 2000US-216251P.

```

```

PR 06-JUL-2000; 2000US-216258P.
PR 06-JUL-2000; 2000US-216768P.
PR 10-JUL-2000; 2000US-217449P.
PR 10-JUL-2000; 2000US-217450P.
PR 10-JUL-2000; 2000US-217660P.
PR 27-JUL-2000; 2000US-221491P.
PR 27-JUL-2000; 2000US-221659P.
PR 27-JUL-2000; 2000US-221670P.
PR 07-AUG-2000; 2000US-223170P.
PR 07-AUG-2000; 2000US-223172P.
PR 07-AUG-2000; 2000US-223460P.
PR 26-OCT-2000; 2000US-244037P.
PR 26-OCT-2000; 2000US-244111P.
PR 26-JUN-2001; 2001US-301217P.
XX
PA (DELT-) DELTAGEN INC.
XX
PI Allen KD, Leviten MW;
XX
DR WPI; 2002-154853/20.
XX
DR N-PSDB; AAD37051.
XX
PT Novel non-human transgenic animal, preferably transgenic mice
PT comprising disruption in target gene, e.g., trypsin gene, useful for
PT identifying an agent that modulates expression or function of target
PT gene.
XX
PS Disclosure; Fig 14; 74pp; English.
XX
CC The present invention relates to non-human transgenic animals preferably
CC transgenic mice comprising disruption in target gene such as trypsin
CC gene. The invention also relates to compositions and methods relating
CC to the characterisation of gene functions. The transgenic animals are
CC useful for identifying an agent that modulates the expression or function
CC of a target. They are useful for identifying an agent that modulates a
CC phenotype associated with a disruption in trypsin genes or limulus
CC clotting factor protease-like genes by administering an agent to the
CC transgenic animal and determining whether the agent modulates the
CC phenotype where the agent has effect on decreased body weight, decreased
CC thymus weight, decreased thymus to body weight ratio, increased pre-pulse
CC inhibition, significant decrease in their response latency to the hot
CC plate test or a decreased response threshold to metrazol. Agents that
CC modulate the expression, function or activity of the target gene are
CC useful for treating a disorder associated with a mutation in trypsin
CC gene or in limulus clotting factor protease-like gene. The transgenic
CC animals are useful for testing the efficacy of proposed genetic and
CC pharmacological therapies for human genetic diseases. They are useful
CC as models for diseases, disorders or conditions associated with
CC phenotypes relating to a disruption in a target and to identify drugs,
CC pharmaceuticals, therapies and interventions which may be effective in
CC treating a disease or other phenotypic characteristics of the animal.
CC The present sequence is calcium-activated neutral protease protein.
CC This sequence is used in the exemplification of the invention.
XX
SQ Sequence 713 AA;

```

```

Query Match 65.6%; Score 2435; DB 23; Length 713;
Best Local Similarity 62.8%; Pred. No. 1.6e-211;
Matches 437; Conservative 126; Mismatches 129; Indels 4; Gaps 3;

```

```

Db      193 KAYKVGSGSYALSGGCTSEAFEDFTGCVTEWYDLQKAPSLYQIILALRSGSLGCSI 252
Oy      243 DITAAOSEALTPOKLIVKGHAYSVTGAEVESNGSLQILIRINPWGEVETGRMNDCP 302
Db      253 NISIRIDLEALTFFKLVKGHAYSVTGAKQVYQOGQVNLIRNRMGEVEMKGPWSDSY 312
Oy      303 SWNTIDPEERERLIRRHEDGEFWMSPFLHYSRLICNLTPTLTSDIYKWKLTMD 362
Db      313 EMNKVDPEREQLRVKMEDGEFWMSPFLHYSRLICNLTPTLTSDIYKWKLTMD 372
Oy      363 GNMWRGSGTAGCCRNYPNTFMNPOYLKLEED--EDEDESGCTFVGLIQKRRQR 420
Db      373 GTWRGSGTAGCCRNYPNTFMNPOYLKLEED--EDEDESGCTFVGLIQKRRQR 432
Oy      421 KMGDMHTIGFQIYEVPEELSGQTNILHSKNFPLTNRARESDPFLNREVLPFKLP 480
Db      433 RFGDMETIGFAVYQVPEELAGQP-VHLKRPFLNARSRAOSEHFINREVSNRIPLPG 491
Oy      481 EYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANL--EEDPISDDIDGVRRLFA 539
Db      492 EYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANL--EEDPISDDIDGVRRLFA 551
Oy      540 QLAEDAEISAFELQTLIRRLAKRQDKSGFSIETCKIWMMLDSGSGKLKERYI 599
Db      552 KLAQDMETISAFELQTLIRRLAKRQDKSGFSIETCKIWMMLDSGSGKLKERYI 611
Oy      600 LMTIIOKYQKTYREIDVDKSGTMSYEMRKALAEAGFPCQLHGVIAVARPADQLIDF 659
Db      612 LMRIRNLYLTFRKFDLDSGSMAYEMRKALAEAGFPCQLHGVIAVARPADQLIDF 671
Oy      660 DNFRVCLVRLTLFKIFKQDLPENTGTIELDLISWL 695
Db      672 DNFRVCLVRLTLFKIFKQDLPENTGTIELDLISWL 707

RESULT 7
AAW41564
ID AAW41564 standard; Protein; 703 AA.
AC AAW41564;
XX
DT 27-APR-1998 (first entry)
XX
De Human calpain.
XX
KW Calpain; human; leukocyte; calcium dependent cysteine protease;
KW screening; activator; inhibitor; treatment; prevention; cancer;
KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;
KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;
KW ischaemic heart disease; atherosclerosis; arthritis.
XX
OS Homo sapiens.
XX
PN EP799892-A2.
XX
PD 08-OCT-1997.
XX
PE 03-APR-1997; 97EP-0105508.
XX
PR 05-APR-1996; 96JP-0083649.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Kawamoto T, Nishi K, Shintani Y;
XX
DR WPI, 1997-482674/45.
XX
DR N-PSDB; AAV04202.
XX
PT Human calpain protein and related DNA - useful for drug screening
PT and treating cancer, stroke, etc.
XX
PS Claim 1; Pages 23-25; 43pp; English.

```

```

XX      The present sequence is calpain, a human leukocyte derived
CC      calcium dependent cysteine protease. Calpain can be used to screen
CC      for compounds that activate or inhibit its proteolytic activity.
CC      Calpain DNA can be used to treat or prevent cancer, cerebral
CC      apoplexy, cerebral infarction, subarachnoid haemorrhage,
CC      Alzheimer's disease, myodystrophy, cataracts, ischaemic heart
CC      disease, atherosclerosis, arthritis or collagen disease.
XX
SQ      Sequence 703 AA;
Query Match 63.8%; Score 2370.5; DB 18; Length 703;
Best Local Similarity 61.5%; Pred. No. 1.1e-205;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;
Oy      1 MAGIAALADREAEGISHERAIKYNQDYALNRECEAGTLFODPSPFPAIPSLAF 60
Db      1 MAQAAGVSGQANATQGLSGNQNALKYLQDPFKLRQCCIDSGVLPFDPPFACPSALGY 60
Oy      61 KELAPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDCWLLAIALSLTNEEI 120
Db      61 KDLGPGSPQIGIIMKRPTELCPSPQITVGATRTDICGALDCWLLAIALSLTNEEI 120
Oy      121 LARVPLNQGFOENYAGIFHFQWQYGEWVYVDRLPRKQDELLFVHABSGSEFSAAL 180
Db      121 LYRVPRDQFOENYAGIFHFQWQYGEWVYVDRLPRKQDELLFVHABSGSEFSAAL 180
Oy      181 LEKAYAKINGCYEALSGATTEGPEDEFTGIAEWELEKPPNLFKIIQALQKSLG 240
Db      181 LEKAYAKINGCYEALSGATTEGPEDEFTGIAEWELEKPPNLFKIIQALQKSLG 240
Oy      241 SIDTSAADSEALTPOKLIVKGHAYSVTGAEVESNGSLQILIRINPWGEVETGRMNDCP 300
Db      241 SIDTSAADSEALTPOKLIVKGHAYSVTGAEVESNGSLQILIRINPWGEVETGRMNDCP 300
Oy      301 CPSWNTIDPEERERLIRRHEDGEFWMSPFLHYSRLICNLTPTLTSDIYKWKLTMD 360
Db      301 APEWNHIDPPRRKEELDKVSDGEFWMSPFLHYSRLICNLTPTLTSDIYKWKLTMD 360
Oy      361 MDGNWRGSGTAGCCRNYPNTFMNPOYLKLEEDDEED--GESGCTFVGLIQKRRQR 418
Db      361 MDGNWRGSGTAGCCRNYPNTFMNPOYLKLEEDDEED--GESGCTFVGLIQKRRQR 418
Oy      419 QKMGEDMTIGFQIYEVPEELSGQTNILHSKNFPLTNRARESDPFLNREVLPFKLP 478
Db      419 QKMGEDMTIGFQIYEVPEELSGQTNILHSKNFPLTNRARESDPFLNREVLPFKLP 478
Oy      479 PGEYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANLAEFDISE--DDIDGVRRL 537
Db      479 PGEYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANLAEFDISE--DDIDGVRRL 537
Oy      481 PGEYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANLAEFDISE--DDIDGVRRL 540
Db      481 PGEYLVPTFEEDNKGDGFCIRVSEKKADYQAVDEIEANLAEFDISE--DDIDGVRRL 540
Oy      538 PAOLAGEDAISAFELQTLIRRLAKRQDKSGFSIETCKIWMMLDSGSGKLKERYI 597
Db      541 FEKLAKDSEITANALIKILNEAFSKRTDIKFGFNINTREMLSLDNGTGLGAVEF 600
Oy      598 YIIMTKIOKYQKTYREIDVDKSGTMSYEMRKALAEAGFPCQLHGVIAVARPADQLIDF 657
Db      601 KTLMLAKIOKLEYIWEVDYHSGTIDAHMRKTLARKAGFTLNSQVQGTIALRYACSLGI 660
Oy      658 DFDNFVACLVRLTLFKIFKQDLPENTGTIELDLISWL 696
Db      661 NFDSEFVACMIRLTLFLSLDLDEKDGWQSLAEWL 699

RESULT 8
AAW41565
ID AAW41565 standard; Protein; 712 AA.
AC AAW41565;
XX
DT 27-APR-1998 (first entry)
XX
De Human calpain.

```

XX Calpain; human; leukocyte; calcium dependent cysteine protease;
 KW screening; activator; inhibitor; treatment; prevention; cancer;
 KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;
 KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;
 KW ischemic heart disease; atherosclerosis; arthritis.
 XX Homo sapiens.
 OS
 PN EP799892-A2.
 XX
 PD 08-OCT-1997.
 XX
 PF 03-APR-1997; 97EP-0105508.
 XX
 PR 05-APR-1996; 96JP-0083649.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Kawamoto T, Nishi K, Shintani Y,
 DR WPI; 1997-482674/45.
 DR N-PSDB; AAV04203.
 XX
 PT Human calpain protein and related DNA - useful for drug screening
 PT and treating cancer, stroke, etc.
 XX
 PS Claim 2; Pages 26-28; 43pp; English.
 CC The present sequence is calpain, a human leukocyte derived
 CC calcium dependent cysteine protease. Calpain can be used to screen
 CC for compounds that activate or inhibit its proteolytic activity.
 CC Calpain DNA can be used to treat or prevent cancer, cerebral
 CC apoplexy, cerebral infarction, subarachnoid haemorrhage,
 CC Alzheimer's disease, myodystrophy, cataracts, ischaemic heart
 CC disease, atherosclerosis, arthritis or collagen disease.
 XX
 SQ Sequence 712 AA;
 Query Match 63.8%; Score 2370.5; DB 18; Length 712;
 Best Local Similarity 61.5%; Pred. No. 1.1e-205;
 Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

QY 1 MAGIAAKLAKREAEAGLGSHERAIKYNODYALNNECLAGTLFQDPSPAPALSGF 60
 DB 10 MAQAAGVSNRQRAATGSGNQNAKYLGGDFKTLRQCCLDSGVLPDFPACPSALGY 69
 QY 61 KEIGPYSKTRGKRWKRPTEICADPOFTIGATRTDICGALGDCWLLAIAISLTNERT 120
 DB 70 KDLGPGSPOTOGIIMKRPTELCPSPQFIVGATRTDICGALGDCWLLAIAISLTNERT 129
 QY 121 LARVPLNOSFOENVGIFHFQFMOYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
 DB 130 LRVVVRDDPQENYVAGIFHFQFMOYGEWVVDRLPTKDELLFVHSEQNEFMSAL 189
 QY 181 LEKAVAKINGCYALSGATTEGFEDFTGIAEYELKPPNLFKIIQALQKSLGCG 240
 DB 190 LEKAVAKINGCYALSGATTEGFEDFTGIAEYELKPPNLFKIIQALQKSLGCG 249
 QY 241 SIDITSAADSEAITFQTLVKGHAYSTGAEEVNSGSLQTLIRINPWEVENTGRMND 300
 DB 250 SIVSSAAEAEAITSQTLVKGHAYSTGAEEVNSGSLQTLIRINPWEVENTGRMND 309
 QY 301 CPSMNTIDPEERELTRRHEDGEFMSFDFLHYSLTICNTLPTDLTSDTYKKYMLTK 360
 DB 310 APFNMHIDPRKEELDKVDEGEFMSFDFLHYSLTICNTLPTDLTSDTYKKYMLTK 369
 QY 361 MDGNWGRGTAGCCRNYPNTFMNAPQYLKLEDEDEDEED--GESGCTPLVGLIQRKRR 418
 DB 370 FNGHMTGRGTAGCCRNYPNTFMNAPQYLKLEDEDEDEED--GESGCTPLVGLIQRKRR 429
 QY 419 QRKMGEDMTIGGIYVEPEELSGQNIHLKSKFELTNRAKRSDFINIREVLNRPKLP 478

DB 430 RKRIGQMGSLIGAVYQVPEKEHSHTDAHLGRDFLAPPSARTSTYVNLREVSGARLP 489
 QY 479 PGEYILVPSSTFPEBNKQDPCIRVFSEKKADYQAVDEIEANLEEFISE-DDIDGVRRL 537
 DB 490 PGEYILVPSSTFPEBNKQDPCIRVFSEKKADYQAVDEIEANLEEFISE-DDIDGVRRL 549
 QY 538 PAQLAGEDAEISAFELQTLIRVLAARODIKSDPSIETFKIWMMLDSDSGSKGLKEF 597
 DB 550 FEKLAKDSEITANALKIILNEAFSKRTDIKFGGFNINTCREWISLSDSGTGLAVER 609
 QY 598 YILMTKIQKQKRYREIDVDRSGTMNSYERKALSEAGFMPQOLHGVYVARADQOLI 657
 DB 610 KTLMLKIQKYLEIYMETDVNHSSTIDAHMERATLARRAGFTLNSQVOQTALRYACSKLGI 669
 QY 658 DFDNFVRLVRLLETLFKIFKQDPENTGTIEDLISMLC 696
 DB 670 NFDSPVACMIRLETLFKLSLDEDDKQGVQSLAEMLC 708

RESULT 9
 AAE14338
 ID AAE14338 standard; Protein; 703 AA.
 XX
 AC AAE14338;
 XX
 DT 07-MAR-2002 (first entry)
 XX
 DE Human protease PRN3-3 protein.
 XX
 KW Human; protease; PRN3-3; tranquilliser; gene therapy; vaccine; allergy;
 KW infection; dermatitis; atherosclerosis; rheumatoid arthritis; hepatitis;
 KW atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;
 KW gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;
 KW epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;
 KW hypertension; neurological disorder; Parkinson's disease; drug screening;
 KW cardiac; cell proliferative disorder; multiple sclerosis; osteoporosis;
 KW diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;
 KW autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;
 KW developmental disorder; reproductive disorder; infertility; diarrhoea;
 KW dementia; acidosis; cataract; gynaecomastia; epilepsy; jaundice.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..508
 FT Domain /note="Calpain catalytic domain"
 FT Domain 45..344
 FT Domain /note="Calpain family cysteine protease domain"
 FT Domain 99..110
 FT Domain /note="Thiol protease Cys motif"
 FT Domain 355..512
 FT Domain /note="Calpain larger subunit domain III"
 XX
 PN WO200183775-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 04-MAY-2001; 2001MO-US14651.
 XX
 PR 04-MAY-2000; 2000US-202082P.
 PR 11-MAY-2000; 2000US-203566P.
 PR 17-MAY-2000; 2000US-205803P.
 PR 25-MAY-2000; 2000US-207477P.
 PR 01-JUN-2000; 2000US-209402P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Deleage AM, Lal P, Hafalia A, Patterson C, Wallia NK, Kearney L,
 PI Tribouley CM, Khan RA, Yao MG, Baughn MR, Azimzai Y, Elliott VS,
 PI Nguyen DB, Gandhi AR, Yang J, Hernandez R, Policky UL, Lu DM,
 PI Reddy R, Yue H, Tang YJ,
 XX WPI; 2002-034518/04.

DR N-PSDB; AAD23844.

XX Novel human proteases and polynucleotides encoding the proteases,
 PT useful for treating, diagnosing or preventing cell proliferative,
 PT cardiovascular, autoimmune/inflammatory, neurological and developmental
 PT disorders -

PS
 XX
 Claim 1; Page 121-123; 151pp; English.

XX The invention relates to human proteases (PRTS-14) and its corresponding
 CC cDNA molecules. Human PRTS and its nucleic acid molecule are useful for
 CC the diagnosis, treatment and prevention of disorders associated with
 CC increased or decreased expression of PRTS. Examples of such disorders
 CC include, cell proliferative disorders (arteriosclerosis, atherosclerosis,
 CC hepatitis, psoriasis and cancers); autoimmune/inflammatory disorders
 CC (AIDS, Addison's disease, allergy, anaemia, asthma, atopic dermatitis,
 CC diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
 CC trauma, Grave's disease, rheumatoid arthritis, ulcerative colitis, and
 CC viral, bacterial, fungal, parasitic, protozoal and helminthic
 CC infections); cardiovascular disorders (myocardial infarction, ischaemic
 CC heart disease and hypertension); neurological disorders (epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, dementia,
 CC Parkinson's disease, stroke, mental disorders including mood, anxiety
 CC and seasonal affective disorder and prion diseases); gastrointestinal
 CC disorders (Crohn's disease, anorexia, nausea, diarrhoea and jaundice);
 CC epithelial disorders (contact dermatitis, eczema, acne vulgaris,
 CC alopecia, scabies, insect bites and urticaria); reproductive disorder
 CC (infertility, disruption of estrous and menstrual cycle and
 CC gynaecomastia); and developmental disorders (renal tubular acidosis,
 CC Cushing's syndrome, seizure disorders, congenital glucocortic acid and cataract).
 CC PRTS DNA is also in useful in gene therapy. PRTS and its immunogenic
 CC fragments are useful for screening libraries of compounds in several drug
 CC screening assays. The present sequence is human protease PRTS-3 protein.

XX
 SQ Sequence 703 AA;

Query Match 63.8%; Score 2367.5; DB 23; Length 703;
 Best Local Similarity 61.5%; Pred. No. 2.1e-205;
 Matches 430; Conservative 122; Mismatches 144; Indels 3; Gaps 2;

QY 1 MAGIAAKIADKREAAEGSGHERAIIKYNODYEARNECEAGTLPDPSPPAIPALGF 60
 DB 1 MAAQAGAGVSRQAAATOGAGSNOMALKYIGDPFKTLRQCLDPSGVLFKDPPEPAPSAIGY 60
 QY KELGPYSKTRGCMKMRPTEICADPOPIIGATRTIDCGALGDCMILAAIATLNEEI 120
 DB 61 KOLGPSPQTOGIIWKRPTEICPSPOPIVGATRTIDCGALGDCMILAAIATLNEEI 120
 QY 121 LARVVPINQSFQENYAGIFHFQFQWQYGEWVEVVDRLPTXGELLFVSAEGSEFWSAL 180
 DB 121 LYRVVPIDQFQENYAGIFHFQFQWQYGEWVEVVDRLPTXGELLFVSAEGSEFWSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEEDFTGIAEWELKKPPNLFKTIQKALQKSLIGC 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEEDFTGIAEWELKKPPNLFKTIQKALQKSLIGC 240
 QY 241 SIDTSAADSEAFIFQKLVKGAHVSATGAEEVNSGOKLIRIRNPGVEYMTGRANDN 300
 DB 241 SIDVSAAEAAITISQKLVKSHAYSVTGVEEVNQGEHEKILRLNPGVEYMSGANSDD 300
 QY 301 CPSWNTIDPEERERLIRRHEDDEFWMSFDFLRHYSRLICNLTPDTLSTDYKKMKLTK 360
 DB 301 APEVNHIDPRKEELDKKVEDGEFWMSSDFVRQSRLEICNLSPDLSSEVHKMNLVL 360
 QY 361 MDGWRKRGSTIGGCRNYPNTFWANPQYILKEEBDEED--GESGCTFLVGLIQKRRR 418
 DB 361 FNGHMTKRGSTIGGCRNYPNTFWANPQYILKEEBDEED--GESGCTFLVGLIQKRRR 420
 QY 419 ORKQMDHTTGFQYVEVEELSGQTNHLKSNFLTRAAERSTPFLNLEVNRFPLP 478
 DB 421 RKRIGQGMSTIGYAVYQVFKELBSHTDAHLGRDFFLAYQPARSTYNNLEVSGRARLP 480
 QY 479 PGEYILVPTPEPNKGDGFCIRVFSEKKADYQAVDEIEANLEEFIDSE-DDIDGVVRL 537

DB 481 PGEYILVPTPEPNKGEFCIRVFSEKKQALRIGVAVGNPYEPHPSEVDQDDQFRL 540
 QY 538 FAQAGDADISAFELQTIARRVLAKRQDIKSPGSIETCKIWMDDSPGSKLGEK 597
 DB 541 FEKLKAGDSITNALKILNLEAPSKRTDKFPGFNINIREMISLSDSGTTLGAVER 600
 QY 598 YIIMTKIQYQKTYREIDVDRSGTMSYEMRKALKEAGFPCQLHQVIVAREPADDQLI 657
 DB 601 KTMIMKIQKLEIYEMEDVNHSGTIDAHMRKTLRRAGFLINSQVQOTIARRVACSLGI 660
 QY 658 DPNFVRCVLRLETLFKIFQDPDENTGTELDLISWLC 696
 DB 661 NFDSPVACMIRLETLFELFLDDEKXQGWQVLSIAEWLC 699

RESULT 10

AAU72884
 ID AAU72884 standard; Protein; 703 AA.
 AC
 XX
 XX AAU72884;
 DT 26-FEB-2002 (first entry)
 DE
 XX Human aspartyl protease partial protein sequence #9.
 XX Human; protease; PCR primer; cytosolic; immunomodulator; cardiant;
 KM vasotropic; antidiarrheal; analgesic; endocrine; nootropic; tranquilizer;
 KM hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
 KM anorectic; antiinflammatory; aspartyl protease; cysteine protease;
 KM metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
 KM lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
 KM immune-related disease; cardiovascular disease; neuronal disease;
 KM migraine; sexual dysfunction; mood disorder; attention disorder;
 KM cognition disorder; hypotension; hypertension; psychotic disorder;
 KM dyskinesia; metabolic disorder; inflammatory disorder.
 XX Homo sapiens.
 OS
 XX
 XX W0200183782-A2.
 PN
 XX 08-NOV-2001.
 PD
 XX
 XX 04-MAY-2001; 2001WO-US14431.
 PF
 XX 04-MAY-2000; 2000US-201879P.
 PR
 XX
 XX (SUGEN-) SUGEN INC.
 PA
 XX
 XX Plowman GD, Whyte D, Sudareanam S, Manning G, Caenepeel S;
 PI Payne V;
 PI WPI; 2002-041502/05.
 DR N-PSDB; AAS97167.
 XX
 XX
 DR Novel protease polypeptide useful for screening for substances that may
 PT be used to treat, e.g., cancers, immune-related diseases,
 PT cardiovascular disease, migraine, pain, psychotic and inflammatory
 PT disorders -

PS
 XX
 Claim 28; Figure 2B; 232pp; English.

XX The invention relates to an isolated, enriched, or purified protease
 CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
 CC screen for substances (S) that may modulate its activity. Administering
 CC S (which modulates protease activity in vitro) may be used to treat a
 CC disease or disorder selected from cancers (e.g., of tissues, of blood or
 CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
 CC brain, ovarian, bladder or kidney), immune-related diseases and
 CC disorders, cardiovascular disease, brain or neuronal-associated diseases
 CC (e.g., central or peripheral nervous system diseases, migraine, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypotension, hypertension, psychotic disorders, neurological

CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (1) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAU72876-AAU72910 represent human
CC protease amino acid sequences of the invention.

XX
Sequence 703 AA;

Query Match 63.8%; Score 2367.5; DB 23; Length 703;
Best Local Similarity 61.5%; Pred. No. 2,1e-205;
Matches 430; Conservative 122; Mismatches 144; Indels 3; Gaps 2;

QY 1 MAGAATLADREAEELGSHERRAIKTLNDYELNRECEAGTLPDPSPFPAISALGF 60
DB 1 MAQAQAVSRORATQGLSGNQNALKYLGDFKTLRQOCIDSGVLFDPPEPAPSPALGY 60
QY 61 KELGYSKTRGRMRKPTTEICADPOFIIIGATRTDICGALGDCWMLAAIASLTANEI 120
DB 61 KDLGFSPTQIGIWKRPTELCPSPQPIVGATRTDICGGLGDCWMLAAIASLTANEI 120
QY 121 LARVPLNQSFOENYAGIFHFQWQYGEWVEVVDRLPTKDELLEFVHSAEGSEFWAL 180
DB 121 LYRVVPEDQDFQENYAGIFHFQWQYGEWVEVVDRLPTKNGQLFLHSEQNEFWAL 180
QY 181 LERAYAKINCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 240
DB 181 LERAYALNCCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 240
QY 241 SIPTSAADSEATFOKLVKGVHAYVNGAEVSENGSLQKLRIRNMGWEWGRNND 300
DB 241 SIPTSAADSEATFOKLVKGVHAYVNGAEVSENGSLQKLRIRNMGWEWGRNND 300
QY 301 CPWNITIDPERERLTRRHDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 360
DB 301 APEWNITIDPERERLTRRHDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 360
QY 361 MDGWRGRSTAGCGRNTPENTFMNPOYLKLEEDDEDEB--GESGTFVLGLIQKRR 418
DB 361 FNGWTRGSTAGCGRNTPENTFMNPOYLKLEEDDEDEB--GESGTFVLGLIQKRR 418
QY 419 ORKMGEMHTIGFQIYVEPELISGQTNHLSKNEFLNRAEREDTINREVYNRFLP 478
DB 419 ORKMGEMHTIGFQIYVEPELISGQTNHLSKNEFLNRAEREDTINREVYNRFLP 478
QY 479 PGEYILVSTFEPKCDGFCIRVSEKKADYQAVDEIEANLEEFIDSE--DDIDGVRRL 537
DB 479 PGEYILVSTFEPKCDGFCIRVSEKKADYQAVDEIEANLEEFIDSE--DDIDGVRRL 537
QY 538 PAQLAGDAISAELOTLIRRLVAKQDIKSDGFSIETCKINVMMLDSGSKLKEF 597
DB 538 PAQLAGDAISAELOTLIRRLVAKQDIKSDGFSIETCKINVMMLDSGSKLKEF 597
QY 598 YIIWTKIOKYOXKIREIDVRSQTMNSYEMRKULEEGFKMPQQLHVIYAFAPDOLI 657
DB 598 YIIWTKIOKYOXKIREIDVRSQTMNSYEMRKULEEGFKMPQQLHVIYAFAPDOLI 657
QY 658 DFDNFVACLVRLETLFKIFKQDPENTGTIELDLSWLC 696
DB 658 DFDNFVACLVRLETLFKIFKQDPENTGTIELDLSWLC 696
QY 661 KTMMLKQKLEIYEWEDYHSGTIDAEHMTAKRKKGFTINSVOQTALRYACSKLGI 660
DB 661 KTMMLKQKLEIYEWEDYHSGTIDAEHMTAKRKKGFTINSVOQTALRYACSKLGI 660

RESULT 11
AAB62152 standard; Protein; 739 AA.

XX AAB62152;
XX
XX 29-MAY-2001 (first entry)
XX
XX Novel human protein (NHP) #1.
XX
XX Novel human protein; NHP; calcium; protease; gene therapy; screening.
XX Homo sapiens.

XX
XX WO200116336-A1.
XX
XX 08-MAR-2001.
XX
XX
XX 01-SEP-2000; 2000WO-US24062.
XX
XX
XX 02-SEP-1999; 99US-0152057.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX
XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2001-235112/24.
XX N-PSDB; AAF57302.
XX
XX Novel human polynucleotides isolated from human testis cDNA library,
XX PT encoded novel human proteins, useful as reagents in screening for
XX PT compounds used for treating mental and biological disorders -
XX
XX Claim 2; Page 27-28; 39pp; English.
XX
XX The invention relates to novel human polynucleotides encoding protein
XX CC (NHP) that share sequence similarity with human calcium dependant
XX CC proteases. Nucleotide constructs encoding functional NHPs are used in
XX CC gene therapy for the modulation of NHP expression. NHP oligonucleotides
XX CC can be used as hybridization probes for screening libraries and assessing
XX CC NHP gene expression patterns. The sequences may also be used as part of
XX CC ribozyme and/or triple helix sequences that are useful for NHP gene
XX CC regulation. Labeled NHP nucleotide probes can be used to screen a human
XX CC genomic library. The NHP nucleotide sequences are also useful in drug
XX CC screening techniques. Nucleotide constructs encoding NHP products can be
XX CC used to genetically engineer host cells to express NHP products in vivo,
XX CC these genetically engineered cells function as bioreactors in the body,
XX CC delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion
XX CC protein to the body. The present sequence represents a NHP.
XX
XX
XX Sequence 739 AA;

Query Match 54.4%; Score 2020; DB 22; Length 739;
Best Local Similarity 52.5%; Pred. No. 7.8e-174;
Matches 366; Conservative 134; Mismatches 193; Indels 4; Gaps 2;

QY 3 GIAPKAKOREAEGLGSHERRAIKYNODEALNRECEAGTLPDPSPFPAISALGF 62
DB 3 GIAPKAKOREAEGLGSHERRAIKYNODEALNRECEAGTLPDPSPFPAISALGF 62
QY 63 LGPYSSKTRGRMRKPTTEICADPOFIIIGATRTDICGALGDCWMLAAIASLTANEI 122
DB 63 LGPYSSKTRGRMRKPTTEICADPOFIIIGATRTDICGALGDCWMLAAIASLTANEI 122
QY 123 RVPPLNQSFOENYAGIFHFQWQYGEWVEVVDRLPTKDELLEFVHSAEGSEFWAL 182
DB 123 RVPPLNQSFOENYAGIFHFQWQYGEWVEVVDRLPTKDELLEFVHSAEGSEFWAL 182
QY 157 RVPVRQGSFPRKNAAGIFHFQWQYGEWVEVVDRLPTKDELLEFVHSAEGSEFWAL 216
DB 157 RVPVRQGSFPRKNAAGIFHFQWQYGEWVEVVDRLPTKDELLEFVHSAEGSEFWAL 216
QY 183 KAYAKINCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 242
DB 183 KAYAKINCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 242
QY 217 KAYAKINCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 276
DB 217 KAYAKINCYELMSGATTEGPEDEFTGIAEMYLEKPPNLFKTIQALQKSLGC 276
QY 243 DITSAADSEATFOKLVKGVHAYVNGAEVSENGSLQKLRIRNMGWEWGRNND 302
DB 243 DITSAADSEATFOKLVKGVHAYVNGAEVSENGSLQKLRIRNMGWEWGRNND 302
QY 277 EVTSDBLESMQKMLVGRHAYSVTGLQDVHGRKMETILIRVANNPGRLEWNAWMSDSAR 336
DB 277 EVTSDBLESMQKMLVGRHAYSVTGLQDVHGRKMETILIRVANNPGRLEWNAWMSDSAR 336
QY 303 SMNTIDPERERLTRRHDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 362
DB 303 SMNTIDPERERLTRRHDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 362
QY 337 EMEVVASDIOQLMHTEDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 396
DB 337 EMEVVASDIOQLMHTEDGEFWMSPDFLRHYSRLICNLPTDITLSTDYKMWLTK 396
QY 363 GNRWRSSTAGCGRNTPENTFMNPOYLKLEEDDEDEDESG--CTFVLGLIQKRR 419
DB 363 GNRWRSSTAGCGRNTPENTFMNPOYLKLEEDDEDEDESG--CTFVLGLIQKRR 419
QY 397 GSWRRGSSAGCGRNHGTFTNPOFKISLPEGDDPEDDAGNVAVCTCLVALAQKWRHA 456
DB 397 GSWRRGSSAGCGRNHGTFTNPOFKISLPEGDDPEDDAGNVAVCTCLVALAQKWRHA 456
QY 420 RKMGEWHTIGFQIYVEPELISGQTNHLSKNEFLNRAEREDTINREVYNRFLP 479
DB 420 RKMGEWHTIGFQIYVEPELISGQTNHLSKNEFLNRAEREDTINREVYNRFLP 479

Db 457 RQGAQLOTTIGFVLYAVPEKQNIQDVHLKKEFTFKYODHGFSEIFTNRSREVSQRLRP 516
Qy 480 GEYILVSTFEPNKGDFCIRVFSEKADYQAVDD-EIENALKEPDISDDIDGVRLLF 538
Db 517 GEYILVSTFEPNKGDFCIRVFSEKADYQAVDD-EIENALKEPDISDDIDGVRLLF 576
Qy 539 AQLGDAEISAFELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGKEFY 598
Db 577 KIVAGEKEIGVVELQRLNRMALFKSFRTKGFQDLACRCIMIMLMDKDGSGKLGLEFK 636
Qy 599 ILMTKIQYQKIVYEDVDRSGTNSYEMRKALBEAGKRCQQLHOYIVARFADQDILID 658
Db 637 ILMTKIKKWMIDIFRECDHSGTNSYEMRLVIEKAGIKLNKKWQVLYAVYADDLLID 696
Qy 659 PDNFVRCIVRETLFKIFKQDPENTGTIELDLISWL 695
Db 697 DPFISCTRLKMTFFPLTMDPKNTGHICLSLEQWL 733

RESULT 12
ID AAM79025 standard; Protein; 702 AA.
AC AAM79025;
XX 06-NOV-2001. (first entry)
DT Human protein SEQ ID NO 1687.
DE Human protein SEQ ID NO 1687.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX WO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dzmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AD, Yang Y, Wehrman T, Goodrich R;
XX WPI: 2001-476283/51.
XX N-PSDB; AAK52158.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 20; Page 4033-4034; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM8020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 702 AA;
Query Match 54.2%; Score 2014; DB 22; Length 702;
Best Local Similarity 52.4%; Pred. No. 2.5e-173;
Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
Qy 4 IAAKLADREABAGLSHERAIIKYNQDYALNECLAEATLPQDPSPAIPALGKEL 63
Db 1 MVAHINNSRLKAGVGQHDNAQNFQNSPEELAAACLRKGLPEDFLPAPESPISGRKOL 60
Qy 64 GPYSKTRGMRWRPPEICADPOFITGATRTDICGALGDCMLAAIASLTINEELAR 123
Db 61 GPNSKNVQNISWOPKDIINNPFIIMDGISPDIICGIGIDCWLAAIGSLTTCPKLLYR 120
Qy 124 VPELNGSPQENYAGIFHFQFQYQGEWVEYVDRLPPTKQELLFVHSAGSEFMSALKE 183
Db 121 VPRGOSFKKNYAGIFHFQIWMQFGQWVNVVDRLPPTKQELLVHSTERSEFMSALKE 180
Qy 184 AYAKINGCYALSGATTEGFEFTGIAEWELAKPPNLFKIIQKALQKSLGCSID 243
Db 181 AYAKLSGYALSGSTMBELGFTGVAQSPQLQRPQVLLMLAKAVRSLSLMGCSIE 240
Qy 244 ITGADSEAITPQKLYKGAHAYSTGABEVESNGSLQKLRIRNPQGEVETGWNDCPS 303
Db 241 VTSDESESEMTDMLVGAHAYSTGLQDVYRKMKETLLRVNRPWRRIENNGAMSARE 300
Qy 304 WNTIDPERBERLRRHEDGSEFMSFSDPLRHYSLREICNLTPTLTSDTYKKKLTMDG 363
Db 301 WEEVASDIQMLHKTEDGSEFMSYDQPLNNTFLLEICNLTPTLTSDTYKKKLTMDG 360
Qy 364 NMRGSTAGGCRNYPNTFMNPOYLITKEEDEDDEDESGS--CTFLVGLQKRRROR 420
Db 361 SMRRGSSAGGCRNHPGTFWNPQKISLPEGDDPEDDAGSNVVCIVLVMKNNRHAR 420
Qy 421 KMGEDMTTIGFIYEVPEELSGQTNHLSKNFPLTNRARSRPTNLREVLNRFKLP 480
Db 421 QCGAQLQTTIGFVLYAVPEKQNIQDVHLKKEFTFKYODHGFSEIFTNRSREVSQRLRP 480
Qy 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDD-EIENALKEPDISDDIDGVRLLF 539
Db 481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDD-EIENALKEPDISDDIDGVRLLF 540
Qy 540 QLAGDAEISAFELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGKEFY 599
Db 541 IVAGEKEIGVVELQRLNRMALFKSFRTKGFQDLACRCIMIMLMDKDGSGKLGLEFK 600
Qy 600 LMTKIQYQKIVYEDVDRSGTNSYEMRKALBEAGKRCQQLHOYIVARFADQDILID 659
Db 601 LMTKIKKWMIDIFRECDHSGTNSYEMRLVIEKAGIKLNKKWQVLYAVYADDLLID 660
Qy 660 PDNFVRCIVRETLFKIFKQDPENTGTIELDLISWL 695
Db 661 DPFISCTRLKMTFFPLTMDPKNTGHICLSLEQWL 696

RESULT 13
ID AAB62154 standard; Protein; 702 AA.
AC AAB62154;
XX 29-MAY-2001 (first entry)
DT Novel human protein (NHP) #3.
DE Novel human protein; NHP; calcium; protease; gene therapy; screening.
KW

XX Homo sapiens.
 OS
 XX
 PN WO200116336-A1.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24062.
 XX
 PR 02-SEP-1999; 99US-0152057.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX
 DR WPI; 2001-235112/24.
 XX
 DR N-PSDB; AAF57304.
 XX
 PT Novel human polynucleotides isolated from human testis cDNA library,
 PT encodes novel human proteins, useful as reagents in screening for
 PT compounds used for treating mental and biological disorders -
 XX
 PS Claim 6; Page 31-33; 39pp; English.
 XX
 CC The invention relates to novel human polynucleotides encoding protein
 CC (NHP) that share sequence similarity with human calcium dependant
 CC proteases. Nucleotide constructs encoding functional NHPs are used in
 CC gene therapy for the modulation of NHP expression. NHP oligonucleotides
 CC can be used as hybridization probes for screening libraries and assessing
 CC NHP gene expression patterns. The sequences may also be used as part of
 CC ribozyme and/or triple helix sequences that are useful for NHP gene
 CC regulation. Labeled NHP nucleotide probes can be used to screen a human
 CC genomic library. The NHP nucleotide sequences are also useful in drug
 CC screening techniques. Nucleotide constructs encoding NHP products can be
 CC used to genetically engineer host cells to express NHP products in vivo,
 CC these genetically engineered cells function as bioreactors in the body,
 CC delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion
 CC protein to the body. The present sequence represents a NHP.
 CC
 SQ Sequence 702 AA;
 Query Match 54.2%; Score 2014; DB 22; Length 702;
 Best Local Similarity 52.4%; Pred. No. 2.5e-173;
 Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
 QY 4 IAAKLADEAAGLGSHERAITYLNODYEALNRECEALGTLPDPSFPAISLGFKEI 63
 DB 1 MVAHINNSRLKAKGVGHDAONFGNOSFEELRAKRGELFEDPLFPAEPSSLAGFKDL 60
 QY 64 GYSSKTRGMRKRPTEICADPOFIIGATRTDICOAGLDCWMLAIAISLTNEILAR 123
 DB 61 GPNKQVONISWQRPKDIINNPFIIMDGISPTDICOGLDWCMLAIAISLTCPKLYR 120
 QY 124 VVPLNGSFQENYAGIFHFQWQYGEVYVVDRLPTFKDCELLFVHSAEGSEFMSALLEK 183
 DB 121 VVPRGOSFKKNYAGIFHFQWQYGEVYVVDRLPTFKDCELLFVHSTERSSEFMSALLEK 180
 QY 184 AYAKINGCYEALSGATTEGFEPTGIAEMWEIKKPPMLFKTIOALKKSILGSGID 243
 DB 181 AYAKLSGYALSGGSTMELGEBDTGVAOSFQIQRPONLRLKRAVVERSSIMGSGIS 240
 QY 244 ITSAADSEALTFOKLKYGAAYSVTGAEVEVNSGLQKLIRIRNPGVEVMTGHWNDGPS 303
 DB 241 VTSDSLEISMTDKMLVGHAYSVTGLQDVHVRGMETLIVRNPWGRIENGMAMSDARE 300
 QY 304 WNTIDPEERERLRRHEDGEFMSFSDPLNHYSLRLCNTLPDTLSYDYKKWLTVMG 363
 DB 301 WEVVASDIOQLHKTEDGEFMSYODPLNNFTLRLCNTLPDTLSGDYSYVHTTTFYEG 360
 QY 364 NMRRGSDAGGGRNYPNTFMMNPQVYLKLEDEDEDESGSG---CTFLVGLTIQKRRRCR 420
 DB 361 SMRRGSSAGGCRNHPGTFMNPQFKISLPRGDEDEDEAGNVVYVCTCLVALMOKNMRHAR 420

QY 421 KMGEDMHTTIGFIYEVPEELISGQTNHLSKNFFLTNRARSDPTINLRREVLRFKLPG 480
 DB 421 QCGAQLQTIGFVLYAVPKERQNIQVHLKKEFFTKYQDHGFSEIFNNSREVSQRLRPG 480
 QY 481 EYTLVSTFERNPDGRCIRVPESEKADYQAVND-ELEANLEFDSIEDIDGVRRLFA 539
 DB 481 EYTLIPSTFEPHDDAPFLRVFTEKHSSEWDEYVAYEQLOEKVSEDDMDDFLHLFK 540
 QY 540 QLAGEDAEISAFELQTLIRVLAKRODIXSDGSFIECTKIMVMDLSDSGSKGLKEFYI 599
 DB 541 IVAGEGKEIGVYELQRLNRMALFKFSFKTKGGLACRCMIMMDKDSGKGLKEFYI 600
 QY 600 LMTKIOKYQIYEIDVDSGTMNSYEMRKALAEAGFKQPCQLHQYTVARFADQGLIDF 659
 DB 601 LMKKLKAMDIREFECQDHSGLTNSYEMRLVIEKAGIKLNKKVQVLYVARADDLLIDF 660
 QY 660 DNFRCLVRLTFLFKTFOKLDPEMTGTEIDLSML 695
 DB 661 DSFISCFRLKTMFTFTFLMDPPNTHICLSLEQWL 696
 RESULT 14
 AAB46595
 ID AAB46595 standard; Protein; 702 AA.
 XX
 AC AAB46595;
 XX
 DT 10-APR-2001 (first entry)
 XX
 DE Human testis-specific calpain CAPN11 protein.
 XX
 KW Calpain; CAPN11; testis-specific protein; fertility regulation;
 KW calcium-dependent cysteine protease; germ cell apoptosis;
 XX testis-specific transcriptional regulator; male fertility disorder.
 OS Homo sapiens.
 OS
 PN DE19928021-A1.
 PN
 PD 21-DEC-2000.
 XX
 PF 18-JUN-1999; 99DE-1028021.
 XX
 PR 18-JUN-1999; 99DE-1028021.
 XX
 PA (BADI) BASF AG.
 XX
 DR WPI; 2001-081733/10.
 DR N-PSDB; AAF25803.
 XX
 PT New nucleic acid encoding testis-specific calpain-11, useful for
 PT identifying specific inhibitors for treatment of fertility disorders -
 XX
 PS Claim 1; Page 8-9; 16pp; German.
 XX
 CC This invention describes a novel nucleic acid (I) encoding human
 CC calpain, designated CAPN11 which has fertility regulating activity.
 CC CAPN11 is a calcium-dependent cysteine protease, expressed in testes and
 CC possibly involved in germ cell apoptosis or regulation of testis-specific
 CC transcriptional regulators. CAPN11 is useful for identifying its specific
 CC inhibitors (A). (A) are used for treating fertility disorders in men or
 CC more generally any condition associated with elevated levels of CAPN11.
 CC
 SQ Sequence 702 AA;
 Query Match 54.2%; Score 2014; DB 22; Length 702;
 Best Local Similarity 52.4%; Pred. No. 2.5e-173;
 Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
 QY 4 IAAKLADEAAGLGSHERAITYLNODYEALNRECEALGTLPDPSFPAISLGFKEI 63
 DB 1 MVAHINNSRLKAKGVGHDAONFGNOSFEELRAKRGELFEDPLFPAEPSSLAGFKDL 60

QY 64 GYSSKTRGMRKRPTEICADPOFIIGGATRTDICOGLADGCMWLLAASLTNBEILAR 123
 DB 61 GPNKKNVONISWQRPKDIINNPLFIMDISPTDICOGLIDGCMWLLAASLTTCPKLXR 120
 QY 124 VVPLNOSQOENYAGIFHFQWQYGEWEVVVDRLPTKOGGLLVHSAEGSEFWSALLEK 183
 DB 121 VVPRGQSFKNYAGIFHFQWQYGEWEVVVDRLPTKNDKLVFHSERSEFWSALLEK 180
 QY 184 AYAKINGCYALSGATGEGFEDFTGIAEWELKPPNLFKIIQKALOGSLGCSID 243
 DB 181 AYAKLSGSEYALSGGTSWEGLEDFTGVASQFQLOPPQNLRLIRKAVESSLSMGCSIE 240
 QY 244 ITSADSEAITFOKLWKHAYSVTGAEEVNSGSLQKILIRNPMGEWEVTGRWNDCPS 303
 DB 241 VTSDEISEMTDKMLVRGHAYSVTGLQDVHYRGKMETLIRVRNPMGRLEMNGAMSDSARE 300
 QY 304 WNTIDPERERLTRRHEDGFWMSFSDFLRHYSRLIECNLTPTDLSSTYKKWKLTKMDG 363
 DB 301 WEEVASDIOQMLHKTEDGFWMSYODFLNFTLIECNLTPTDLSGDKSYMHTTFYEG 360
 QY 364 NMRGSGTAGGCRNYPNTFPMNPOYLKLEEDDEDEDESG---CTPLVGLIQKRRRQR 420
 DB 361 SMRSGSSAGGCRNHPGTWNPQFKISLPEGDDEDDAEGVNVVCTCLVALMOKMWHAR 420
 QY 421 KMGEDMHTIGFQIYEVEEELSQOTNHLKSNFPLTNRARSDFINLREVLNRFKLPFG 480
 DB 421 QQGAQLOTIGFVLAVPFQONIDVHLKKEPFTKYODHGSSEIFFTNSREVSQRLRPPG 480
 QY 481 EYILVPSTFEPNKGDFCIRVFSEKADYQAVDD-EIEANIEEDFISDDIDGVRRLFA 539
 DB 481 EYIILPSTFEPHARDFLRVFTEGHSESWELDEVNAEQLEEVSDMDQOFLHLPK 540
 QY 540 QLAGEDATISAFELQTLIRVLAKRODKSDGFSIETKIIWMDMDSGSGKLGLKERYI 599
 DB 541 IVABEGKEIGYELORLNRAIKFKSKTGFGIDACRCMINDKSGSKGLLEPKI 600
 QY 600 LMTKIQKQKIYREIDVRSSTNMSYEMRKALBEAGFPMQOLHVOIVAREPADOLIIDF 659
 DB 601 LMKLTKKMWIDIFRECDQDHSGTSLNSYEMRLVIERKAGIKLNKMWQVAVAYADDLIIIDF 660
 QY 660 DNEPRCLVRLTLPKIFKQDLPENTGTIELDLISL 695
 DB 661 DSFISCFRLKTMFTFFLTMDPKATGHICLSLEQWL 696

RESULT 15
 AAU72887
 ID AAU72887 standard; Protein, 702 AA.
 AC XX
 AAU72887;
 XX XX
 DT 26-FEB-2002 (first entry)
 XX XX
 DE Human aspartyl protease partial protein sequence #12.
 XX XX
 KW Human; protease; PCR primer; cytosolic; immunomodulator; cardiant;
 KW vasotropic; antilipolytic; analgesic; endocrine; nootropic; tranquiliser;
 KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
 KW anorectic; antinflammatory; aspartyl protease; cysteine protease;
 KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
 KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
 KW immune-related disease; cardiovascular disease; neuronal disease;
 KW migraine; sexual dysfunction; mood disorder; attention disorder;
 KW cognition disorder; hypotension; hypertension; psychotic disorder;
 KW dyskinestia; metabolic disorder; inflammatory disorder.
 XX XX
 OS Homo sapiens.
 XX XX
 PN MO2001.83782-A2.
 XX XX
 PD 08-NOV-2001.
 XX XX
 PF 04-MAY-2001; 2001WO-US14431.

XX 04-MAY-2000; 2000US-201879P.
 PR (SUGC-) SUGEN INC.
 XX PA
 XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S,
 PI Payne V,
 XX
 DR WPI; 2002-041502/05.
 DR N-PSDB; AAS971170.
 XX
 PT Novel protease polypeptide useful for screening for substances that may
 PT be used to treat, e.g., cancers, immune-related diseases,
 PT cardiovascular disease, migraine, pain, psychotic and inflammatory
 PT disorders -
 XX
 PS Claim 28; Figure 2C; 232pp; English.
 XX
 CC The invention relates to an isolated, enriched, or purified protease
 CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
 CC screen for substances (S) that may modulate its activity. Administering
 CC S (which modulates protease activity in vitro) may be used to treat a
 CC disease or disorder selected from cancers (e.g., of tissues, of blood or
 CC hematopoietic origin, of the breast, colon, lung, prostate, cervical,
 CC brain, ovarian, bladder or kidney), immune-related diseases and
 CC disorders, cardiovascular disease, brain or neuronal-associated diseases
 CC (e.g., central or peripheral nervous system diseases, migraine, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypertension, psychotic disorders, neurological
 CC disorders and dyskinesias), metabolic disorders and inflammatory
 CC disorders. (I) may also be useful as a diagnostic tool for a disease or
 CC disorder such as those above. AAU72876-AAU72910 represent human
 CC protease amino acid sequences of the invention.
 CC
 SQ Sequence 702 AA;
 Query Match 54.2%; Score 2014; DB 23; Length 702;
 Best Local Similarity 52.4%; Pred. No. 2,5e-173;
 Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
 QY 4 IAAKLADREAAAGLSGHERAIKYNODYALNNECLGAGTLQODSFPALPAGKEL 63
 DB 1 MVAHINNSRLKAGVGQHDNAQNFQNSFEELRAACLRKGLPEDFLPPEPSLQKDL 60
 QY 64 GYSSKTRGMRKRPTEICADPOFIIGGATRTDICOGLADGCMWLLAASLTNBEILAR 123
 DB 61 GPNKKNVONISWQRPKDIINNPLFIMDISPTDICOGLIDGCMWLLAASLTTCPKLXR 120
 QY 124 VVPLNOSQOENYAGIFHFQWQYGEWEVVVDRLPTKOGGLLVHSAEGSEFWSALLEK 183
 DB 121 VVPRGQSFKNYAGIFHFQWQYGEWEVVVDRLPTKNDKLVFHSERSEFWSALLEK 180
 QY 184 AYAKINGCYALSGATGEGFEDFTGIAEWELKPPNLFKIIQKALOGSLGCSID 243
 DB 181 AYAKLSGSEYALSGGTSWEGLEDFTGVASQFQLOPPQNLRLIRKAVESSLSMGCSIE 240
 QY 244 ITSADSEAITFOKLWKHAYSVTGAEEVNSGSLQKILIRNPMGEWEVTGRWNDCPS 303
 DB 241 VTSDEISEMTDKMLVRGHAYSVTGLQDVHYRGKMETLIRVRNPMGRLEMNGAMSDSARE 300
 QY 304 WNTIDPERERLTRRHEDGFWMSFSDFLRHYSRLIECNLTPTDLSSTYKKWKLTKMDG 363
 DB 301 WEEVASDIOQMLHKTEDGFWMSYODFLNFTLIECNLTPTDLSGDKSYMHTTFYEG 360
 QY 364 NMRGSGTAGGCRNYPNTFPMNPOYLKLEEDDEDEDESG---CTPLVGLIQKRRRQR 420
 DB 361 SMRSGSSAGGCRNHPGTWNPQFKISLPEGDDEDDAEGVNVVCTCLVALMOKMWHAR 420
 QY 421 KMGEDMHTIGFQIYEVEEELSQOTNHLKSNFPLTNRARSDFINLREVLNRFKLPFG 480
 DB 421 QQGAQLOTIGFVLAVPFQONIDVHLKKEPFTKYODHGSSEIFFTNSREVSQRLRPPG 480
 QY 481 EYILVPSTFEPNKGDFCIRVFSEKADYQAVDD-EIEANIEEDFISDDIDGVRRLFA 539

```

Db 481 EYIITPSTFEPHRDADFLRLRVTEKHSBMSWELDEVNTAEOLOEKEVSEDDMODFLHLFX 540
QY 540 QLAGEDAEISAFELQOTILRRVLAERODIKSDGFSIETCKIWDMLDSGSGKLGKEFYI 599
Db 541 IVAGEGKEIGVYELQRLNRMALFKSGFKTGFGIDACRCMINIMDKDSSGKLGLEFKI 600
QY 600 LMTXIOXYOKIYREIDVDRSGTMNSYEMRKALBEAGFPMPCOLHOVIVARPADDLIIDF 659
Db 601 LMKLKKWMDIFRECQDQHSSTNSYEMRLVIERKAGIKLNKVMQVLVARYADDDLIIIDF 660
QY 660 DNFVRCIVRLLETLEFKIKOLDPENTGTIELDLISWL 695
Db 661 DSFISCFRLRLKTMFTFPLTMDPKNTGHCISLBQWL 696

```

Search completed: July 24, 2003, 12:51:44
 Job time : 91.9793 secs

